**The Dissertation Committee for Cheryl Annette Browne Certifies that this is the approved version of the following dissertation:**

# Preschoolers' Explanations for Intentional and Unintentional Behavior

**Committee:**

Jacqui D. Woolley, Supervisor

Randy L. Diehl

Frederick Kronz

Judith H. Langlois

Ann Repp

# Preschoolers' Explanations for Intentional and Unintentional Behavior

**by**

**Cheryl Annette Browne, B.A., M.A.**

**Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**December, 2004**

# Dedication

For my family.

## Acknowledgements

So many people who contributed to this culmination of my graduate studies I can't possibly name them all. As the nature of causality is that every member of the causal chain somehow affects the eventual outcome, in truth everyone I've known for the last several years has played a role in the nature and indeed existence of this dissertation.

Nevertheless, I thank all those whose positive influence is foremost in my mind. First, I acknowledge the role played by my academic advisor, Jacqueline Woolley, whose academic mentorship, friendship and role as confidante were absolutely vital to my ability to undertake and complete this project. No less important was the role played by my family, particularly by my husband, Matt, and my parents. Most appreciated is the emotional support they offered, particularly during exceedingly difficult periods during the past several years.

Finally, I sincerely thank the dedicated undergradtuate student assistants , without whose assistance this project would undoubtedly have been enormously difficult to complete. I especially thank Chelsea Cearly and Young Song, whose hard work and enthusiasm kept me going through the whirlwind of the final few months before graduation.

I'm quite sure that my decision to attend the University of Texas was one of the best I've ever made.

**Preschoolers' Explanations for Intentional and Unintentional Behavior**

Publication No._____

Cheryl Annette Browne, Ph.D.

The University of Texas at Austin, 2004

Supervisor: Jacqui D. Woolley

In this dissertation I begin by discussing and evaluating various models for how people causally explain behavior in their everyday discourse. Using logical argument and empirical evidence I endorse a folk psychological model proposing that people explain intentional and unintentional actions differently. When a behavior is seen as being intentional it is usually explained by "reasons," which specify the actor's beliefs and/or desires that led to the intention to act. On the other hand, when a behavior is perceived to be unintentional it is predominantly explained by "causes," which make reference to non-psychological forces on the actor that bring about action directly, without being mediated by intention.

In two studies I investigated 4- to 6-year-olds' understanding of the relation between the intentionality of an action and the type of explanation used to explain it. Experiment 1 consisted of trials in which children were told about two protagonists performing the same action; one was explained with a reason, and the other with a cause. Children indicated which protagonist performed the act on purpose. Experiment 2 was the reverse; each trial consisted of one story about a protagonist who performed an

intentional or unintentional action. Children chose between a reason and a cause explanation for the act.

Overall, children performed significantly above chance level for both studies, but when age groups were considered separately only the two older groups' performance exceeded chance. This finding suggests that children begin to recognize the relation between action type and explanation type around the beginning of their sixth year, which is consistent with past studies showing related developments at that age.

Performance on Experiment 1 was somewhat better than on Experiment 2, and only Experiment 1 showed an age effect. It is argued that these findings, combined with the fact that in Experiment 2 the "intentional" and the "unintentional" items were uncorrelated, suggest that two separate domains of knowledge – about the mind and about physical objects, with their separate characteristic modes of causation – become appropriated for the crucial task of explaining human behavior.

# Table of Contents

## List of Figures

## INTRODUCTION

"Explanation is a river that flows through human life"

(Wilson & Keil, 2000, p. 87)

Imagine you're tired and want to rest, so you sit down in a chair. Now imagine that a strong wind knocks you into a chair that happens to be behind you, leaving you in a sitting position. Despite appearing similar in some ways, these two actions are quintessentially different. In the first case, you had an intention to sit down before you actually did it. You believed that sitting would allow you to rest. You may have thought about which chair you wanted to sit in before choosing one. You may have deliberately looked behind you as you were sitting in the chair. In the latter case, you had no particular thoughts about sitting before ending up in a sitting position. You weren't thinking about resting or about chairs. You did not take any deliberate action as you were sitting; it just happened. After ending up in a sitting position you were surprised, rather than feeling like you had accomplished your goal.

These two types of behavior – intentional and unintentional behavior – do not only feel different. They are also treated differently by others. For example, the law recognizes this difference in its treatment of murder versus manslaughter, and intentional acts of helping and aggression are more likely to be reciprocated than are the same acts that are unintentional (Malle, 1999).

This difference between intentional and unintentional behavior is a key aspect of a model of how we explain behavior, cited by a few attribution theorists and several philosophers (e.g., Buss, 1979; Heider, 1958; Malle, 1997, 1999; Malle, Moses, and

Baldwin, 2001; Searle, 1983). According to this model, we see intentional behavior as being generated by intentions, which in turn are generated by reasons (usually beliefs and desires), whereas we see unintentional behavior as being generated by causes, which usually generate behavior in some physical way. The model, therefore, can be represented as follows: reasons → intention → intentional behavior; causes → unintentional behavior.

The distinction between how reasons and causes generate behavior has been compared to Aristotle's views of causality. Aristotle proposed four types of causality – material, formal, final, and efficient – the latter two of which correspond to reasons and causes. Aristotle's "final cause," which is "its function or the good it serves" is similar to a reason (in the example above, the "function" of sitting would be *to rest*). In turn, his "efficient causality" is analogous to what we are calling "cause" (i.e., the wind that knocked us over) (Lloyd, 1995).

The model described above, which is grounded in ancient philosophy, provides the framework for this paper. It has been tested and verified with adults (Malle, 1997, 1999; White, 1991) – that is, it has shown to describe how people actually explain behavior – but has never, in its entirety, been tested with children. Therefore, although much research has been conducted that is relevant to certain parts of the model, we do not know whether children see reasons as generating intentional behavior and causes as generating unintentional behavior.

First, I will expound on the model[1] of how we explain intentional and unintentional behavior in more detail. Then I will present some different versions of the

[1] Henceforth I will call this model "Malle's model," even though several authors have proposed similar models.

model, as well as several criticisms that have been levied against it, arguing that Malle's model can be supported both empirically and conceptually. Next, I review the research relevant to the question of whether, and to what extent, children's folk reasoning can be described by the model.

## MODEL OF HOW WE EXPLAIN INTENTIONAL AND UNINTENTIONAL BEHAVIOR

### Description

In a classic book, which is often said to have established the entire field of attribution theory, Heider (1958) developed the cause/reason dichotomy, in which explanations for intentional and unintentional behavior are seen as logically distinct. However, Heider did not use the labels of "reasons and causes;" rather, he used the terms "personal causality" and "impersonal causality." His model, like the one endorsed in this paper, can be represented as "reasons → intentions → intentional action" and "causes → unintentional behavior." Though similar, his model was also different from Malle's model (1997, 1999) in important ways, which we will see later.

Heider's idea of separating intentional from unintentional action in terms of how we explain behavior lay dormant for many years, until the seventies and the early eighties, when Buss (1978, 1979) as well as Locke and Pennington (1982) revived it. In arguing for the distinction between intentional and unintentional behavior, and between causes and reasons, Buss wrote that "causes and reasons are logically distinct categories for explaining different aspects of behavior." Specifically, "causes are that which brings about a change," whereas "reasons are that for which a change is brought about (e.g., goals, purposes, etc.)" (1978, p. 1311). Buss went on to argue that unintentional (or irrational) behavior tends to be explained by causes, whereas intentional behavior tends to be explained by reasons.

Locke and Pennington (1982) similarly explained the difference between reasons and causes: "Whereas causes explain behavior by showing it to be the automatic or inevitable outcome of a certain complex of conditions in accordance with the relevant

laws of nature, reasons explain behavior by showing it to be what any rational agent would do, given the relevant beliefs and desires" (p. 213). They added that an agent always knows what his reasons are, but is not necessarily aware of the causes of his behavior.

Buss's (1978, 1979) and Locke and Pennington's (1982) arguments were not sufficient to make a significant impact on the field, however. Attribution theorists have tended to, and still do, focus more on the distinction between "person explanations" and "situation explanations," or "external" versus "internal" explanations for behavior (e.g., Jones & Davis, 1965; Kelley, 1967; Shaver, 1975). For example, "she sneezed because of the pollen" would be a situation explanation, because it refers to something in her environment but not herself, whereas "she reached up because she wanted to open the cupboard" would be a person explanation because it refers to her and her desires. Although this dichotomy is similar to the reason/cause distinction, it does not map onto it directly.

Recently, Malle (1997, 1999; Malle, et al., 2001) appealed to other attribution theorists to bring back the cause/reason distinction for explaining unintentional versus intentional behavior. As discussed before, his basic view of our folk-explanations of behavior is that causes directly influence behavior in some mechanical way without an intervening intention (e.g., sadness causes crying; sunshine causes happiness) (Malle, et al., 2001), whereas reasons are beliefs and desires that generate behavior via intention (e.g., He wanted to go to the store, and believed that driving there would help him get to the store, therefore he intended to drive). Thus, we explain intentional behavior with the beliefs and desires in light of which the intention was formed, and we explain unintentional behavior with generating factors that are not reasons and do not lead to an intention.

According to Malle (1997, 1999), although reasons usually appear as beliefs or desires, the words "want" and "think" don't have to actually appear; these can be implicit, or "unmarked," as in "I ran to catch the bus" instead of "I ran because I wanted to catch the bus." Unmarked beliefs sometimes seem like they're not beliefs at all, as in, "She went to the beach because it was sunny." In this case it is her *belief* (i.e., her knowledge of the fact) that it was sunny that caused her to form an intention to go to the beach, not the sunniness itself. This example is a belief-reason, as opposed to a desire-reason; Malle claimed that desire-reasons are the "primary motives of action" and are more common than belief reasons.

Malle (1997, 1999; Malle, et al., 2001) clarified that while causes and reasons are the two main modes of explanation, there are also two minor modes: enabling factors, which clarify how the intention led to the action, such as "he ran a mile because he's in good shape," and causal history of reasons, which clarifies the factors that led to the reasons. Causal history of reasons is used when the observer is either ignorant of the actor's mental states that led the actor to act, as in, "I don't know why he said that, *weird people do these kinds of things*" (p. 276), or when the observer wishes to express a whole class of reasons, as in "I go to the supermarket every day because I have three kids" (p. 277). It is not having three kids that actually causes the speaker to go to the supermarket; rather, having three kids instigates other, more immediate reasons (e.g., "I'd like to get some diapers, and going to the store is the best way of achieving that aim"), perhaps different ones each time he goes.

It was said earlier that the cause/reason distinction is not in vogue in attribution theory. Indeed, to my knowledge only one attribution theorist's views were sufficiently similar to Malle's (1997; 1999) to include in the discussion of his theory. Nevertheless, views of commonsense psychology similar to his have been alive and well within

philosophy for some years. I now turn to the views of these philosophers, along with the one relevant attribution theorist, White (1991), pointing out where they agree with the model and where they propose ideas that refine the model.

**Philosophers' views**

Like attribution theorists Malle (1997, 1999; Malle, et al., 2001) and White (1991), philosophers Lennon (1990) and Dretske (1988) maintained that intentional actions are explained in terms of purpose or reasons, whereas unintentional acts are explained by antecedent causes: "Crucially," Lennon explained, with intentional acts "we are seeking to understand the point or purpose of the activity from the agent's point of view," whereas "we do not expect [an unintentional action's] explanation to differ fundamentally from that offered for other natural phenomena – for example, one billiard ball hitting another and causing it to move" (p. 16). Similarly, White wrote, "the basic distinction . . . is between behavior seen as conscious, voluntary, and intentional and explained in terms of reasons, and behavior seen as unconscious and deterministic, and explained in terms of causes other than reasons" (p. 261).

Woodfield (1976), who expounded on the explanation of intentional behavior, explicated the meaning of the common form of explanation "X did Y in order to Z": "Mary wants to go shopping in order to buy a present" means "Mary wants to go shopping because she wants to buy a present and believes that going shopping will contribute to buying a present" (p. 217). Here we have both a desire (wants to buy a present) and belief (that going shopping will help her buy the present). We can think of this as Mary intending to go shopping, and the intention is formed in light of the relevant desire and belief.

Probably the philosopher most known for his views on the intentionality of behavior is Searle (1983). Although quite complex, his basic philosophy of behavior was

very similar to that of Malle (1997, 1999). He agreed that intentional behavior is fundamentally different from unintentional behavior, and that an intention to do X results from a desire for A, as well as a belief that action X will result in A. Searle used a slightly different lexicon from Malle, however. He called unintentional acts simply "bodily movements" but intentional acts "actions." To illustrate the difference between bodily movements and actions he contrasted two ways of generating a hand movement: stimulating an electrode planted in the motor cortex of one hemisphere of the brain, and voluntarily moving one's hand. "There is an obvious phenomenal difference between the case where one moves one's hand and the case where one observes it move independently of one's intentions…Such concepts as 'trying', 'succeeding', and 'failing' apply to it in ways that they do not apply to the experiences the patient has when he simply observes his hand moving" (p. 90).

Searle (1983) clarified several aspects of the basic model of explanation and intentionality. One is that in order for someone to do something intentionally, the intention actually has to generate the action; the behavior cannot come about in some other way. For example, Searle argued, if someone intended to kill his uncle, then drove so recklessly that he inadvertently hit a pedestrian who happened to be his uncle, his killing his uncle would not be considered intentional. Thus, it is not enough to have an intention that matches the outcome; the two have to be causally related.

Searle (1983) also introduced the idea that the intention does not have to be prior to the action; it can be an "intention-in-action." When we switch gears while driving, for example, we may not have thought beforehand, "I'm going to switch gears." Nevertheless, switching gears is intentional because we have the intention in some sense *while* the action is happening.

Another of Searle's (1983) ideas about the issue was that sometimes the action is a non-action, as when one intends to hold still or to refrain from doing something. Refraining from performing an action is not itself a *behavior* or an *action*, nonetheless it is something we can think about doing beforehand and that can result from our beliefs, desires, and intentions. In this sense, the absence of behavior is intentional.

Davidson (1963) did not deal with unintentional behavior, but he discussed intentional behavior in detail. His model was basically the same as Malle's (1997, 1999, Malle, et al., 2001), but he added that sometimes we perform actions that are intentional under one description but not under another. To illustrate his point, he explains that when we turn on the light we are also illuminating the room, flipping the switch, and inadvertently alerting a prowler that we are home. Here we are not doing four things, but only one, of which four descriptions have been given. The behavior is intentional under some descriptions, but unintentional under others.

Malle's (1997, 1999) model of our explanation of behavior can also be refined by considering that reasons consist of both a desire and a belief, but one can often be inferred from the other. For example, I can explain that I left this morning at 8:30 because I wanted to get to work on time. One can infer from this statement of desire that I also believed that leaving at that time would allow me to get to work on time. The opposite is also true – if I say that I left for work at 8:30 because I thought that doing so would allow me to be on time for work, it can be inferred that I also want to get to work on time, that I believe such an outcome to be desirable.

It should be mentioned that like most psychological models Malle's (1997, 1999) model of how we explain behavior should not be considered absolute; it should be seen as capturing how we *usually* explain behavior, not how behavior *must* be explained, neither in the sense of natural law nor logical necessity. In fact, it is easy to come up with

exceptions to the model. One example of unintentional behavior being described by reasons is when we explain unintentional bodily functions in terms of their purpose for the body as a whole. That is, one could say that human beings sweat in order to cool themselves. (Bodily functions are often explained in terms of reasons because the concept of evolution tends to be misconstrued as perhaps a semi-intentional force that designs species to maximize survival.)

Also, causes can sometimes be mental states, but are not the mental states in light of which the intention is formed and the action is performed. Consider the example, "She put salt in her coffee because she thought it was sugar." Here, the explanation for the mistake contains a mental state. However, assume that pouring salt was the action, having sweet coffee was the desired outcome, and jar A contains salt. She did not *decide* to pour salt *in light of* her false belief that it was sugar. In fact, she could not have done anything in light of her false belief because she wasn't aware that her belief was false.

**How do we identify intention?**

Next I turn briefly to the question of how we know whether behavior is intentional or unintentional. How do adults make that distinction? In addition to Malle (1997) this question has been addressed by a few authors working in the "theory of mind" field in psychology, which tackles issues regarding children's understanding of both their own and others' minds.

There are many ways of distinguishing between intentional and unintentional behavior. For example, if the behavior itself is witnessed, one can often tell by determining whether or not the actor is looking at or aware of what she is doing (Malle, 1997). Also, one can judge whether the outcome is likely given the actor's skill in the particular behavior. If the actor is considered to be unable to perform a particular behavior, the behavior and the outcome will be seen to be unintentional, or by chance.

For example, if someone is goofing around, flings a dart towards the board and hits the bullseye, the outcome is not considered to be intentional, but rather driven by chance. Most importantly, intentional actions have a certain "look" to them showing the actor to be careful, directed, and coordinated in his or her actions. If the action is merely described, one can often make the distinction according to the verb used. As pointed out by Astington (1999), intentionality is deeply entrenched in our language. For instance, some verbs inherently imply intentionality and others non-intentionality. The verb "die" is non-agentive and implies non-intentionality whereas "murder" is agentive and implies intentionality.

As discussed by Wellman and Phillips (2001), our reactions toward the result of an action can also be a clue to its intentionality. When we do something intentionally we look happy at the success of an action, and if we fail we might say something like "oops" or "darn." These examples imply that the person was trying to achieve some action. Even 14- to 18-month-old infants show an understanding of vocal cues to intentionality or non-intentionality by imitating actions accompanied by "there!" but not actions accompanied by "oops" (Carpenter, Akhtar, & Tomasello, 2000). Also, when someone does something unintentionally they might look surprised or make an utterance denoting surprise, like "oh."

**Testing the model with adults**

The reason I am endorsing Malle's (1997, 1999) model for how we explain behavior is not simply that similar views are popular with philosophers. Rather, the model has been tested empirically in adults, by Malle himself as well as by White (1991). The basic model has been verified, but as yet many of the details and refinements have not been addressed empirically.

In two straightforward studies, Malle (1999) demonstrated that people tend to explain unintentional behavior with mere causes, and intentional behavior with reasons. In the first study he was interested in showing that when people encounter a reason explanation for a behavior, they will tend to see the behavior as intentional, but when they encounter a cause explanation for the same behavior, they will see it as unintentional. He presented to participants two behaviors that were ambiguous as to intentionality (driving above the speed limit and interrupting one's mother) and explanations that were either reason explanations (e.g., she wanted to get to the store before six, or she knew the store closed at six) or cause explanations (e.g., she wasn't paying attention to the speedometer, or the speedometer didn't work). Participants rated how intentional they thought each behavior was on a nine-point scale. As hypothesized, participants rated the behavior as much more intentional when it was explained with a reason explanation than when it was explained with a cause explanation.

The second study was the reverse of the first one in that Malle (1999) tested the hypothesis that when people are asked to explain an unintentional behavior, they will offer cause explanations, but when they are asked to explain an intentional behavior, they will offer reason explanations. Participants were presented with twenty behaviors, and they rated the intentionality of the behaviors and gave explanations for them. The explanations were then coded for being causes or reasons. For example, "the speedometer broke" was coded as a cause of speeding, whereas "she wanted to get to the store" was coded as a reason. The results were again in favor of the model – intentionality ratings correlated strongly with reason explanations. That is, behaviors that tended to be seen as intentional also tended to be explained with reason explanations, and vice versa.

In a study very similar to Malle's (1999) Study 1, White (1991) had participants rate behaviors with reason explanations and cause explanations attached. They rated the

behaviors as being "conscious" or "unconscious" as well as "intentional" or "unintentional." The behaviors explained with reasons were judged to be conscious and intentional, whereas the behaviors explained by causes were seen as being unconscious and unintentional.

In a series of four studies, Malle (1997) addressed the intentional aspect of the model – its reasons → intentions → behavior component. In Study 1, adults rated behaviors as being intentional or unintentional. Some participants were asked to make their intentionality judgments from the actor's perspective, and others were asked to make the judgments from the observer's perspective. Both groups had access to the same information. Agreement was high both across and within conditions. This study provides evidence that actors and observers agree on intentionality and therefore that intentionality is a shared concept, not dependent on the perspective from which people are judging or on who is doing the judging.

Study 2 (Malle, 1997) explored adults' notions of intentionality. Participants explained in their own words what it means to do something intentionally. Ninety-six percent (of the 87% who did not give mere synonyms for intention) mentioned beliefs, desires, or intention, or some combination of these, as well as the awareness of doing the behavior while it is being done.

In Study 3, Malle (1997) tested people's opinions about whether skill in a particular behavior is necessary in order to judge a behavior to be intentional. One of the two scenarios in the studies involved a man who had been practicing throwing a penny such that it always lands on tails. However, he hasn't been able to succeed in getting tails at more than chance level (skill absent). The test question was, "Do you think X achieved outcome Y intentionally?" The results indicated that skill was indeed necessary. People tended to see unskilled performance of a behavior requiring skill, such as biasing

a coin throw or succeeding in a particularly difficult combination shot, as happening by chance. It should be noted that skill is only relevant to certain types of behaviors; for most behaviors, such as going to the store, skill is not an issue.

Finally, in Study 4 (Malle, 1997) the self-awareness of one's actions was manipulated, and adults were asked whether the behaviors were intentional. That is, for some of the behaviors that were presented the actors were aware of what they were doing while they were doing it, whereas for other behaviors they were not. Awareness was deemed to be necessary for intentionality of behavior. For example, Frank's bumping into George's car was seen to be accidental if Frank didn't know he had bumped into the car.

**Arguments against Malle's model**

Although views similar to Malle's (1997; 1999) model have appeared within philosophy for many years, they have not escaped criticism. Also, several versions of the model have been proposed that differ from Malle's in important ways. In the next section I will address arguments against the model (and against similar models that have been proposed over the years), and discuss the differing versions of the model.

One criticism that has been levied is that the cause/reason distinction is not a true dichotomy, or not a useful one. Kruglanski (1979), for example, argued that because cause and reason both bring about change in an "antecedent-consequent" relation (p. 1448) and do so in a lawful and predictable way, the two types of explanation are not sufficiently distinct to make the dichotomy useful. It is true that an intention occurs before the action, just as a cause does, even though the content of an intention refers to the future (e.g., "I want to eat *to feel sated"*). His claim that both cause and reason bring about their changes in a lawful way refers to the fact that when people intend to do things they usually do them. Rarely does an obstacle impede the intention → behavior relation.

Thus, in Kruglanski's (1979) terms, "there is nothing unique about classifying explanations into causes and reasons." (p. 1453) Reasons are simply one kind of generating factor, and causes are generating factors that are non-reasons. According to him, the reason-cause distinction may be compared to the possible distinctions between any of the infinite explanatory categories and their complements, such as mechanical versus monmechanical explanations, biological versus nonbiological explanations, and so forth.

Kruglanski's claim (1979) that reasons and causes are both antecedent-consequent generating factors is true. However, this fact does not lead to his conclusion. First, *any* generating factor by definition comes before the effect. Further, that does not mean the distinction is arbitrary. The most important difference is that reasons refer to mental states, goals, and purposes, whereas causes usually refer to some physical aspect of the situation without referring to mental states. They refer to fundamentally different kinds of causality (mental versus physical).

In addition, cause and reason, when used as explanations, demarcate intentional from unintentional behavior, which as argued earlier is a fundamental distinction. Moreover, as pointed out by several authors (Buss, 1979; Harre & Secord, 1972; Taylor, 1964), cause and reason are logically distinct. Reasons have "representational content" (I want *to take a walk*), but causes do not. In other words, we do not just desire and believe; there is content to our desires and beliefs. We desire and believe *something* (Schueler, 2001). Another way in which causes and reasons are logically distinct is that one could, in principle, explain the exact physical mechanism whereby a cause generates a behavior (e.g., if one knew enough about biology one could describe precisely how hitting a knee with a hammer causes the leg to move); in contrast, we have no idea what intentions are, neurologically, nor how they lead to behavior, and it is likely that it is inexplicable, at

least with the sciences we have today. Thus, cause and reason are conceptually distinct; one cannot be deduced from the other (Taylor, 1964).

A similar argument was proposed by Harvey and Tucker (1979). Their quarrel was not with the cause/reason distinction but with the intentional/unintentional distinction. They contended that the distinction was ambiguous because there are behaviors that one cannot say are truly intentional or unintentional, such as manifestations of emotionality, like smiling. Along the same lines, they suggested that it would be possible for a behavior to be started unconsciously, but then finished intentionally.

Harvey and Tucker (1979) were saying, in essence, that the distinction between intentional and unintentional behavior is invalid because there are in-between cases that do not quite fit in one category or the other, that cannot be said to be definitely intentional nor unintentional. However, the existence of in-between cases does not render the distinction invalid; there are many useful distinctions in psychology for which there are in-between cases. For example, in learning theory we distinguish between operant and classical conditioning, yet discriminant learning could be described as involving either operant or classical conditioning, depending on which aspect is being emphasized. Also, Shwitzgebel (1999) expounded on the idea that there are in-between cases of belief, cases in which one would not be correct to say either that one definitely believes something or definitely does not believe it. Clearly, this does not mean that belief is an invalid or useless concept. By the same token, in-between cases do not make the intentional/unintentional distinction invalid or useless.

A further argument relates to the idea that reasons are always inside the person, whereas causes are outside the person. When Heider (1958) wrote his seminal book on the explanation of behavior, he not only made the cause/reason distinction, but he argued

that the inside/outside the person dichotomy mapped onto it. It is the inside/outside distinction that was retained in mainstream attribution theory; in fact, it often seems to be confused with the reason/cause distinction. For example, Lalljee, Watson, and White (1983), in a study in which children gave explanations for intentional and unintentional behavior, first wrote about reasons/causes as differentiating intentional from unintentional behavior, then seemed to equate reasons with "person" (inside) explanations and causes with "situational" (outside the person) explanations. The explanations were coded into person/situational factors.

Although I agree that reasons are in an important sense internal to the person, as they are mental states, I contend that the person/situational factor distinction should be discarded and replaced by the reason/cause dichotomy, untainted by concern with whether explanations are personal or situational. As Malle (1999) has argued, the person/situational dichotomy ignores unintentional behavior that is caused by person factors. For example, if I shiver and explain the behavior by saying "because I am cold" the explanation would be considered a "person factor." Shivering is unintentional, however, and being cold is not a reason, it is a cause. For this reason, although the person/situation distinction is often treated as though it demarcates intentional from unintentional behavior, it actually tends to obliterate the intentional/unintentional distinction when it is used in research. As argued by Malle (2001), it "fails to acknowledge the central role that people's concept of intentionality plays in shaping their explanations of behavior" (p. 265).

Further, when authors code explanations in terms of internal/external features they tend to make the distinction based on surface features of the explanations. For example, "She went because her mom was there" would be coded as external (because it mentions her mother and not herself) whereas "she went because *she thought* her mom

was there" would be coded as internal. The explanations are paraphrases of each other, and should both be coded as reasons (because they both are about her belief), yet in one study they were coded by attribution theorists as different types of explanations (Malle, 2001). Another example was provided by White (1991) in making the same point: "I want to make a lot of money" and "chemistry is a high-paying field" (p. 260) as explanations for choosing chemistry as a career path would be coded as internal and external respectively, even though they are both reasons and both explain an intentional behavior.

In fact, White (1991) showed empirically that the intentional/unintentional distinction does not map onto the internal/external distinction. In a study in which participants rated explanations for intentionality as well as whether they were internal or external, the two distinctions did not correlate. Thus, the two distinctions should not be considered interchangeable, as they often are.

A final difference between Malle's model and contrasting versions of it pertains to the question of whether actors and observers differ in their explanations. Buss (1978) addressed this issue in his attempt to bring the cause/reason distinction back into the field. His claim was that intentional behavior is explained by the actor using reasons, whereas it is explained by the observer using either reasons or causes. Similarly, Locke & Pennington (1982) suggested that for intentional behavior, actors tend to give reasons, whereas observers tend to give causes. Both Buss and Locke and Pennington maintained that unintentional behavior is explained by causes. The difference is that while Locke and Pennington believed that observers tend to give causes for intentional behavior, Buss' view was that observers can give either reasons *or* causes for intentional behavior.

Buss's (1978, 1979) and Locke and Pennington's (1982) claim that unintentional behavior is explained by causes is uncontroversial. However, as shown by Malle (1999;

Malle et al., 2001), observers and actors do not actually differ in their explanations. I agree that intentional actions are sometimes explained by what seem to be "causes;" however, many of these "causes" are what Malle (1997) termed "causal history of reasons" (Davidson, 1963, discusses this as well). For example, if Mary robs a bank, you might cite her poor background in your explanation of her behavior, as opposed to her reasons for doing so (e.g., wanting money and believing that robbing a bank is a good way to get money). However, Mary's poor background shaped her beliefs and desires, which led to her behavior. Thus, what you are really doing is giving the causal history of her beliefs and desires, not explaining the action itself. Also, this type of explanation is not limited to observers; to use the previous example, Mary herself could cite her own poor background to explain her behavior.

## RESEARCH WITH CHILDREN

Much research has been conducted both on children's explanations and on their understanding of intention. However, the two areas rarely meet in research, and never has the question of whether children explain intentional behavior with reasons and unintentional behavior with causes been addressed in a valid way. Nonetheless, examining the research on these separate areas provide the background and context of the present experiments.

I limit my review to studies that include young children, as research suggests that important milestones in both explanation/causality and the understanding of intention are reached between 3 andd 5 years of age.

### Preschoolers' explanations

First I will review the literature on preschoolers' explanations to give a sense of what these children's explanations, in general, are like. Then I will explore the more pertinent issue of how preschoolers explain human behavior.

One study by Callanan and Oakes (1992) shows the kinds of explanations children receive at home. They conducted a diary study in which for two weeks mothers wrote down the "why" and "how" questions asked by their preschool children as well as the explanations that ensued in response. Explanations were coded into "prior cause" and "consequence." These two types of explanation correspond roughly to causes and reasons; for example, "I moved my leg because someone pushed it" would be coded as "prior cause" in this coding scheme, and as "cause" in Malle's (1997, 1999) model. On the other hand, "I moved my leg so I could kick someone" would be a "consequence" in this coding scheme, and a reason in Malle's model. It was found that children were receiving both types of explanations, but that the prior cause explanations were more

frequent than the consequence explanations. The fact that children receive both types of explanation at home bodes well for their ability to supply their own explanations.

In a naturalistic study looking at preschoolers' own explanations (McCabe & Peterson, 2001), their spontaneous speech was analyzed for correct uses of the causal connectives "because" and "so." Surprisingly, preschoolers (4- to 6-year-olds) made no more errors (such as reversing cause and effect, as in "I am cold because I want to go inside") than did adults, suggesting that children are adept at the causal form (X because Y, Y so X, etc.). Also, participants of all ages used the causal connectives more for psychological causality than for physical causality. Perhaps because other people are so important in our lives children learn to talk about them with sophistication, using the causal words "because" and "so", before they learn this skill with objects.

In a similar but longitudinal study (Hood & Bloom, 1979), naturalistic speech was collected from 2- to 3-year-old children for a period of 17 months, and was analyzed for various aspects of causal utterances (ones that include "why" or "how" in a causal way, that are responses to causal questions, or that otherwise imply a causal relationship between clauses). Included in the findings was a substantial increase over time in causal utterances in general, suggesting that as children reach preschool age they are becoming more interested in and adept at causal explanation. In addition, the earliest responses to adults' causal questions were unintelligible, but the frequency of causally interpretable responses increased markedly over the 17-month period. Further, the children's first causal statements did not include the connectives "because" or "so," but the proportion of statements with connectives increased as the children reached preschool age. In addition, as in the McCabe and Peterson (2001) study errors in the use of "because" and "so" were rare. Also consistent with that study, children tended to talk about psychological

causality, using explanations that included intentions and motives, more than about physical causality.

Following Piaget's research on artificialism, Gelman and Kremer (1991) explored preschoolers' explanations for the origins of natural phenomena, such as the ocean and dirt, as well as the origins of artifacts. Four- to 5-year-olds were asked if people were responsible for the phenomena, then were asked "How?" or "Why not?". Responses were scored as "naturalistic" (coming from nature, from a living thing, from an animate natural process such as growth, or from an inanimate natural process) or "nonnaturalistic" (all other responses). Results showed that when the preschoolers were asked, "How?" or "Why not?" they gave naturalistic responses only for natural kinds, never for artifacts. For our purposes, the important finding is that the children responded appropriately to the causal queries, varying their explanations according to the phenomenon in question.

Two conclusions can be drawn from these studies. First, preschoolers are interested in causality and explanation, as evidenced by the frequency of their causal utterances; they are starting to use causal explanation in their own speech even before the beginning of their preschool years. Second, the findings suggest that by four years of age children can reliably answer open-ended "why" questions with causal responses (although the quality or correctness of the explanations might be lacking). Note that responding causally is not the only way children could answer a causal question – they could merely repeat the question or respond with a non-causal aspect of the situation, or say nothing at all. But they do not do this. Responding causally is quite an achievement, as it has only been 3 years since these children began talking at all.

Not only can they respond to open-ended "why" questions, but preschoolers can also choose the most sensible explanation out of several choices. In four studies (Springer & Keil, 1991), preschoolers chose among different types of explanation for why flowers

are pink, why some dogs are brown, and why some cans are red. For color acquisition, preschoolers rejected "silly" explanations such as "a little man came along with a paintbrush, opened the seed carefully, and painted it pink" (p. 770) and "the mother flower wanted her baby to be pink just like her" (p.772) for the biological kinds, and "the machine wanted the can to be red just like the other cans" (p. 779) for the artifacts, in favor of biological explanations for the biological kinds and human intention explanations for the artifacts.

In addition to showing that preschoolers can distinguish between reasonable and unreasonable mechanisms for color acquisition, this study suggests that preschoolers can systematically choose appropriate explanations for "why" questions from among several options. More broadly, these studies on children's explanations suggest that children use causal language in their natural speech, and that this increases from 2 years old to at least 4, as well as that preschoolers can choose reasonable causal explanations out of several choices.

**Preschoolers' explanations of behavior**

With the exception of Kalish (1998), none of the following studies deal with the cause/reason distinction, or Malle's (1997, 1999) model of the explanation of human behavior, *per se*. However, several address certain parts of the model and give us a general idea of how preschoolers explain other people's behavior.

Probably the most basic question with regard to young children's explanations of behavior concerns their ability to explain actions with reference to mental states when the actor's desire is clear and the action is obviously intentional. In a classic study, Bartsch and Wellman (1989) showed that 3-and 4-year-olds indeed possess this ability. The authors invoked explanations for intended action (e.g., "Jane is looking under the piano for her kitten. Why do you think Jane is doing that?") (p. 949) from 3- and 4-year-olds as

well as adults. The authors coded the responses for psychological reasoning (beliefs and desires). Both children and adults appealed to beliefs and desires in their explanations for behavior (as opposed to physical or behaviorist responses, etc.). This study speaks to the intentional part of the model, in which beliefs and desires are seen to play a causal role in behavior. Although it does not address young children's knowledge of the cause/reason distinction or their knowledge about intention as a mental state, it shows that even 3-year-olds know that when a person wants something and performs an action that could lead to achieving it, the behavior is explained in terms of the desire for the outcome or the belief that the action will lead to the desired outcome.

This finding, coupled with research cited above showing that children employ physical causality to explain non-human events, raises the question of how children decide which type of causality to use in their explanations of events in general (e.g., human behavior as well as natural or mechanical events). Proponents of the most prominent proposal, the "entity" view of explanation, claim that young children think of each kind of entity as having its own type of causality (e.g., Carey & Spelke, 1994). For instance, Carey and Spelke proposed that children regard people solely in terms of goals, feelings, and other psychological states, and that they regard inanimate objects in terms of non-intentional, physical causality. Whereas preschoolers' treatment of inanimate objects and events has not been addressed systematically, the studies discussed below show that children can be flexible in their explanations of human behavior. Young children seem to recognize that the same type of entity (people, in this case) can be reasoned about in different ways; in addition to intentional beings people are also bodies moving through space and obeying the causal laws of other physical objects.

Before presenting evidence suggesting that children's explanations of human behavior is flexible, I should point out that his flexibility is necessary in order to explain

behavior in terms of both reasons and causes. In fact some awareness that non-mental forces influence behavior is probably necessary for much younger children as well, though not necessarily in the form of verbal explanations. If we thought of every movement or action exclusively in terms of purposes, beliefs, and desires, our experience of the social world would be radically different. Attempting to explain or predict behaviors such as falling down, breaking a favorite glass, or snoring would be perplexing if they were believed to be purposeful. Such a belief would lead to the view that people are irrational beings with obscure, conflicting desires and odd ways of obtaining desired outcomes[2]. For this reason it is unlikely that children possess a pure "entity" view at any point in development, and it is not surprising that researchers tend to find explanatory flexibility in regard to human behavior.

In one study providing evidence against the entity view, Hickling and Wellman (2001) coded children's speech from CHILDES database of natural child language for causal utterances. Their main goal was to determine what kinds of things young children (2.5- to 5-years-old) explain and how children link explanation types to entity types. The explanations were divided into 6 categories: physical, psychological, social-conventional, biological, behavioral and magical. The results showed that children produced more physical explanations than any other kind, but that people were most often explained.

[2] Our social experience would be analogous to that of Funes, a character in a story by Borges, who as the result of falling off a horse acquired the unfortunate ability to process and remember events in minute detail and lost the crucial ability to parse these details into discrete, meaningful events. He viewed each "action," such as the twitching of a lip or the plunge of a dagger into an abdomen, as equally salient and memorable. In Funes's words, the sum of his experiences and memories is "like a garbage disposal" (Borges, 1967).

(The authors do not specify what aspects of people the children were explaining; behavior? Appearance? General tendencies?) The authors concluded that young children "explained people" not just in psychological terms (e.g., "the man wants to eat"), but also in physical, biological, and social-conventional terms (e.g., a biological explanation might be, "the man is fat because he ate food").

This study shows that 1) children are interested in and knowledgeable about causality, 2) in particular, they're interested in people's behavior, and 3) they explain human behavior not just in psychological terms, but in other terms as well: "Children explain intentional, goal-directed action in psychological terms but more passive, object-like movements in physical terms" (Hickling & Wellman, 2001; p. 677). However, it is difficult to interpret Hickling and Wellman's data specifically in terms of the question of whether children explain intentional behavior using reasons and unintentional behavior using causes, because the data are not coded in these terms. It might seem that the authors' judgments of psychological versus physical explanations would map onto reasons and causes, and in some cases they do. For example, "I talking very quiet because I don't want somebody to wake me up" (p. 678) was coded as psychological and would be coded as a reason in the reason/cause framework. "I pushed it because I got knocked down" (p. 678) was coded as physical and would be a cause in the reasons/causes framework.

However, it is not always the case that their physical/psychological dichotomy maps onto the cause/reason distinction. For example, the utterance, "I can't step on the binoculars because they will break apart" (p. 679) was coded as physical (the justification being that binoculars breaking focuses on a mechanical force), whereas it would be considered a reason under the reason/cause framework because it implies a belief (that

they will break apart) in light of which the decision not to step on the binoculars was formed.

Other research questioning the entity view of explanation includes a series of studies by Schult and Wellman (1997). The purpose of their research was to determine whether 3- and 4-year-olds explain different sorts of human actions (caused by psychological states, physical forces, or biological processes) in different and appropriate ways. In the first study the children explained voluntary actions, mistakes, physically-caused and biologically-caused behaviors. The 4-year-olds, but not the 3-year-olds, gave psychological explanations for the voluntary actions and mistakes and gave physical and biological explanations for the other two types.

This sounds very much like it tests whether children's explanations conform to Malle's (1997, 1999) model of explaining behavior. However, although the authors acknowledged that mistakes, physically-caused and biologically-caused behaviors were all unintended, they did not analyze the data in a way that takes this into account (i.e., they did not combine these categories). Also, it may seem that the psychological/physical dichotomy in this study maps onto the reasons/causes dichotomy in our framework, but this is not the case. Several of the examples cited in the article were coded in a way that belies this mapping. For example, for one mistake story in which a character accidentally pours ketchup on his ice cream, a child responded, "She just didn't know it was ketchup" (p. 297). This was coded as a psychological explanation, yet according to Malle's model the false belief directly causes her to pour ketchup without an intervening intention to do so.

Although this response makes references to beliefs and desires, the character's lack of knowledge would be a cause under the reasons/causes framework because the character did not desire to pour ketchup. The character's false belief led directly to her

pouring ketchup without an intervening intention to do so. Further, pouring ketchup could easily be construed as the outcome instead of the action. According to this perspective reaching for a bottle (which she incorrectly took for chocolate sauce) would be the action and it would be intentional. Thus, the data support the claim that young preschoolers can shift from one type of explanation to another depending on the action, but the findings do not support a conclusion that children easily and appropriately shift from reason to cause explanations depending on the intentionality of action.

In a pilot study, the authors had tried telling stories and asking 3-year-olds for explanations, but children's responses were mostly incomprehensible or "don't know," or the children changed the story so that the outcome matched the desire. Neither this nor any of the studies on children's general explanations mentioned so far have shown that 3-year-olds can express explanations of given behaviors, so it remains possible that as Schult (2002) concludes, "3-year-olds in general are not good explainers" (Schult, 2002, p.1739).

Because of the difficulty eliciting causal responses from 3-year-olds, in a second study the authors (Schult & Wellman, 1997) elicited yes/no judgments by 3- and 4-year-olds as to whether people who want to perform certain impossible actions can perform those actions (e.g., float in the air, stay awake forever). The children responded appropriately, which was interpreted by the authors as indicating that even 3-year-olds' explanations refute the entity view by showing that people's actions are constrained by physical and biological forces. However, the understanding that people cannot violate physical laws is only one small component of the ability to explain actual actions with regard to physical forces.

A study by Kalish (1998) comes the closest of all the studies discussed so far to testing whether preschoolers conform to the reason/cause model of how we explain

behavior. However, as we see later, the coding scheme does not match the causes/reasons distinction, and the questions are about why people *do not* do things rather than why people *do* things. The aim of the study was to investigate children's understanding of the distinction between social rules (voluntary conformity) and physical laws (automatic conformity). The author presented 3- and 4-year-old children with stories about people who wanted to perform actions that would violate each type of law. They were queried about whether the characters could perform the actions, then were asked, "How?" or "Why not?" Responses were coded in terms of reasons versus causes. Causes were generally physical limitations (e.g., "he's not tall enough to touch the ceiling") whereas reasons referred to social injunctions (e.g., "his parents won't let him") or to adverse consequences of performing the action (e.g., "he would ruin his shoes") (p. 709).

Kalish found that many of the children's justifications were coded as neutral (neither cause nor reason), but of the justifications that were coded as either reasons or causes more of them were cause than reason explanations for the physical laws, and for social rules the 4-year-olds (but not the 3-year-olds) gave more reason than cause explanations. These data suggest that children use both causes and reasons in their explanations in a systematic way. Because following physical laws does not require intention, whereas following social laws usually does, these data seem quite suggestive that children adhere to our model of how we explain behavior.

Although it is encouraging that a developmental psychologist is making uses of the reason/cause dichotomy, there are a few limitations in this study that somewhat undermine the conclusions. First, the 3-year-olds gave few explanations that were reasons or causes – for the physical stories fewer than half were coded as one or the other. This makes the data somewhat difficult to interpret. More importantly, the coding did not quite follow the traditional cause/reason framework. For example, possible negative

consequences were coded as reasons (e.g. "his shoes would get wet [if he took a bath with his clothes on]") (p. 712). However, this should only count as a reason if the boy was aware of this consequence and decided not to take a bath in light of this knowledge. There is no indication that this line of reasoning took place.

Similarly, statements like "his parents wouldn't let him" (p. 709) were coded as reasons, whereas it is possible that the children were thinking that his parents were physically stopping him from performing the action, as opposed to providing a law that he intentionally followed. If the parents were physically stopping him, this would be a cause, not a reason. In addition, statements of impossibility ("it is impossible for a boy to grow a beard") (p. 709) were coded as causes, whereas impossibility could be meant in a psychological way. For instance, I might consider it impossible to bungee jump, not because gravity is suspended, but because of my lack of courage, in light of which my intention is formed.

Finally, the main question in the study, regarding why someone cannot perform a desired but physically impossible action, is importantly different from the question of why someone *did* perform an action. An appropriate explanation for the latter question is necessarily a causal one. To explain why someone did something is to establish the factor(s) that generated the action. On the other hand, the relationship between the failure of an impossible (but desired and attempted) action and its explanation is not causal. Several arguments support this claim.

Consider the case of the child in Kalish's study (1998) who responded to the question, "Why can't this boy [who was attempting to touch the ceiling] touch the ceiling?" by saying that the boy wasn't tall enough. This answer is perfectly sensible, but not being tall enough and being unable to touch the ceiling are not events or actions but rather ongoing states of affairs, and one state of affairs (not being tall enough) does not

actually generate, or cause, another state of affairs (i.e., his inability to touch the ceiling). When we think of a causal chain, it is difficult or impossible to think of one state of affairs generating another state of affairs, and that state of affairs generating another state of affairs, etc., without any definable intervening events or actions occurring.

Relatedly, probably the most basic criterion for our considering a relationship to be causal, one that even preschoolers utilize in their causal judgments (e.g., Woolley, Browne, & Boerger, 2004), is that the cause must always precede the effect. However, in this case the boy's not being tall enough does not *precede* his inability to touch the ceiling; both are states of affairs that exist simultaneously.

A final argument is that the relationship between not being tall enough and being unable to touch the ceiling is not causal but rather logical. As touching the ceiling is possible if and only if the boy is above a certain height, and it's the case that the boy is not above that height, then it can be deduced that he cannot touch the ceiling. Causal relationships, on the other hand, do not have this logical relationship; an effect can be predicted, but cannot directly be deduced from a cause.

Because of these limitations, the conclusion that preschoolers can explain human behavior appropriately with reasons and causes should be considered tentative. Nevertheless, these studies corroborate the claim that 4-year-olds do not reason about people's behavior solely in psychological terms. They realize that people's intentions to act are sometimes superceded by physical forces.

In summary, preschoolers older than 3 are able to explain or choose explanations for (non-human) events in a causal way, as well as explain human behavior. Further, under certain circumstances they can explain behavior not solely in terms of psychology, but in other ways as well. That is, they seem to grasp that human beings are physical objects in the world that are subject to physical forces in addition to psychological agents

with beliefs, desires and goals. These are important prerequisites for being able to understand both causes and reasons in explanations of human behavior.

**Children's understanding of the relations among the components of intention**

It is well known that up to age three children do not have a mature notion of intention. They cannot, for example, distinguish between intention and desires (e.g., Astington, 2001; Wellman & Phillips, 2001). According to Astington (2001) toddlers have a general notion that conflates desire and intention. Supporting this idea is the fact that children do not use such terms as *intend to, on purpose,* or *mean to* until their preschool years. But as early as 1 ½ years they talk about goals and desires, primarily with the term *want* (Wellman & Phillips, 2001). In fact, Bartsch and Wellman (1995) analyzed 200,000 utterances of children aged 2-5, and found over 5000 genuine references to desire (e.g., "want").

What about preschoolers? Do they have a mature notion of "intention" that includes the correct relations among its components? The findings of research on their ideas about intention are quite complex and sometimes contradictory.

Two studies by Astington speak to this question. The author interprets the results as suggesting that 3- to 5-year-olds can distinguish action from desire and outcome, but that they may equate intention with action. In one study with 3- to 5-year-olds (Astington, 1999), children were told a series of stories in which each of two characters wanted something but only one actually acted to get it. The other actor obtained the object through some stroke of luck, although he was not trying to get it. The following is an example:

> "Bert is at the apple farm. He wants an apple off the tree. He can't reach it. He climbs on a box. He still can't reach it….Ernie is at the apple farm too. He wants an apple off the tree. It's too high up. He doesn't know what to do. Then the apple

> falls off the tree, onto the ground beside Ernie…Which guy was trying to get an apple off the tree?" (p. 308)

Even the youngest children were able to pick out the one that was "trying to" obtain the desired object, despite the fact that he did not obtain it. These results suggest that 3- to 5-year-olds recognize "trying" when it is associated with an obvious action, and can distinguish it from both desiring and obtaining. It is also possible, however, that 3-year-olds simply associate "trying" with "not getting" and conversely "getting" with "not trying," and therefore chose the picture associated with the character who did not end up getting the object. As in many of the studies cited here, these limitations mitigate the apparent conclusion, which in this case is that children as young as 3 conceptually differentiate among desire, action, and outcome.

In another study by Astington (1991), 3-year-olds could not distinguish between action and intending/preparing, but 5-year-olds could. In this study, 3- to 5-year-olds were shown pictures of two characters – one who was performing an action and one who was preparing to (intended to) perform the action. The point of the study was to see whether children would choose different pictures for a question about actions ("Who's painting?") and one about intentions ("Who's gonna/thinks he'll/wants to/would like to paint?") (p. 167). A majority of the 5-year-olds, but few 3-year-olds, distinguished between action and intention by choosing the action picture for the action question and the "preparing to" picture for the intention question. Here 3-year-olds showed, by choosing the action picture for the intention question, that they associate intention not with the mental state that it is, but with action.

Data from this study (Astington, 1991) suggest that 5-year-olds, but not 3-year-olds, understand that intentions/desires differ from their subsequent actions. Astington argues that this change comes at around 4 years of age, the same time that children come

to see beliefs as representational. She suggests that there is a general change in children's understanding of mental states at about 4 years of age that underlies both of these understandings.

Supporting the idea of 4-year-olds coming to see intentions differently from 3-year-olds is another set of studies by Feinfield, et al. (1999), in which 3-year-olds could not distinguish between intention or desire and outcome, but 4-year-olds could. In the first study, the children heard stories in which the characters' intentions differed from both the desire and the outcome. The characters wanted to go to location A, but decided (intended) to go to location B because their mothers told them to go there. However, they inadvertently ended up at location A, which was where they wanted to go. Children were asked where the characters thought they were going to go, where they tried to go, and where they wanted to go. If they answered the intention questions (where he thought he was going and where he tried to go) according to desire, they would be incorrect. Similarly, if they saw matches between the desire and outcome as indicating that the act of going to location A was intentional (using the "matching rule"), they would also be wrong. To be correct, they must understand that intentions are mental states that do not necessarily match desires or outcomes.

In this study (Fienfeld, et al., 1999) the 4-year-olds performed significantly better on the "try" questions, the "think" questions, and nearly significantly better on the "want" questions than did the 3-year-olds. The 3-year-olds performed significantly worse than chance on the "try" and "think" questions (those asking about intention), whereas the 4-year-olds performed significantly better than chance on those questions. Both age-groups performed significantly better than chance on the "like" question (asking about desire).

In a second, similar study in which the desire differs from the intention (Fienfeld, et al.), 3- and 4-year-old children heard stories in which a character's mother sends him

to find an object in a location, and there the character finds a different, more desirable object than the one he was originally looking for. Children were asked what the character was trying to get and what he thought he was going to get. Again, the 4-year-olds performed significantly better than the 3-year-olds. The older children were significantly above chance levels on the two intention questions ("What does he think he'll do," "what is he trying to do"), but the 3-year-olds were only slightly above chance on the "try" question and were at chance on the "think" question. Thus, the 4-year-olds are again are shown to be more sophisticated than 3-year-olds in their concept of intention.

Still another study pitting desire against intention is Schult's (2002) Experiment 2, in which 3-, 4-, and 5-year-olds played a game involving tossing bean bags into buckets, some of which contained a prize. In each trial they chose beforehand which bucket to aim for but they did not know which buckets contained prizes. After each trial they were asked what they were trying to do, as well what they had wanted (it was assumed that they wanted the prize, although which bucket they were aiming for had nothing to do with whether there was a prize in the bucket). All age groups were able to report their intention correctly when they hit the bucket at which they were aiming. However, whereas the 4- and 5-year olds scored 100% on these questions even when they did not hit the aimed-for bucket, the 3-year-olds were not able to separate their intentions (to hit the bucket they were told to aim for) from their desires (to get a prize) when they missed the bucket they were aiming for and got a prize anyway. They tended to say they were trying to hit the bucket with the prize all along, although they had no way of knowing that the prize was in that bucket.

Here the 3-year-olds seem to have been using a desire-outcome matching strategy (the "matching rule"). That is, when their desire to get a prize matched what actually happened, they inferred that getting a prize was what they were intending to do. What

these children did *not* understand is that the fulfillment of one's desired outcome does not by itself imply that a purposeful act has occurred. Rather, one's desired outcome can come about in various ways, only a subset of which suggests that his or her action was purposeful. Specifically, for one's action to be considered intentional one has to intend to perform the action, actually perform it, and possess the skill and luck to act successfully from the actor's point of view.

Four- and 5-year olds, on the other hand, showed a more mature understanding of intention. They did not assume that when a bucket with a prize was hit that bucket was the intended one. Rather, they limited their claims that they had done what they were trying to do to instances in which they hit the bucket for which they'd been told to aim, thereby demonstrating their ability to distinguish their desires from their intentions and from the outcome.

Based on the previous several studies, 4-year-olds do seem to have a more mature idea of intention than do 3-year-olds. However, they have a long way to go before their notion of intention is adult-like. In a study by Baird and Moses (2001), 4-year-olds had trouble attributing different desires to two people performing the same action, whereas 5-year-olds had no trouble with this. The authors told 4- and 5-year-old children stories in which two characters performed the same action but had different desires (running home for dinner versus running to get healthy and strong). They were then asked what each character tried to do (e.g., get somewhere fast versus get some exercise). Although 5-year-olds could attribute different intentions to the actors with the same action, 4-year-olds tended to attribute the same intention to the actors. As noted by the authors, 4-year-olds' pattern of response suggests that their understanding of intention may still be inextricably tied to action. Thus, although Astington (1991) hypothesizes that children learn to separate intentions from actions (or, as she puts it, see intentions

"representationally") at about 4 years of age, these data (Baird & Moses, 2001) show that 4-year-olds have not quite reached that stage yet.

In sum, what do these studies indicate about the ability of preschoolers to coordinate the relations among the various components of intention? First consider 3-year-olds. Research suggests that that they are just beginning to be able to relate some of the components to each other. Although they have trouble distinguishing intention from action (Astington, 1991), desire from intention (Fienfeld, 1999), and desire and intention from outcome (Schult, 2002)**,** they can use the "matching rule" to infer the intentionality of an action based on whether the desire matches the outcome. The matching rule was a liability in some of the more difficult studies, in which the desire did not match the intention or outcome, but everyday experience suggests that in the majority of cases it is an effective strategy that even adults use when some of the relevant information is not available.

Regarding older preschoolers, some of the research just reviewed shows that 4-year-olds' understanding of the relations among the various components of intentional action is better than that of the 3-year-olds, but is still far from adult-like. Their capabilities include distinguishing between intentions and desires (Fienfeld, et al., 1999) and even between desires, intentions, and outcomes (Schult, 2002), which is impressive given the similarity between desire and intention. On the other hand, Baird and Moses (2001) showed that they are not able to attribute different intentions to people who are performing the same action, even when the intentions are clearly spelled out. They also have deficits in judging whether intentions have been acted upon, tending to say that someone has acted intentionally whenever the desired outcome is fulfilled (Schult, 2002).

Not many of the studies about the relations among the components of intentionality included 5-year-olds, and still fewer address relations that might distinguish

them from 4-year-olds. Astington (1991) showed that 5-year-olds could distinguish between action and preparing/intending, though 3-year-olds could not. It is not clear, however, how 4-year-olds would perform in relation to 3- and 5-year-olds. In addition, they could attribute different intentions to people who were performing the same action (Baird & Moses, 2001). Finally, like 4-year-olds they are able to separate intentions from desires and outcomes (Schult, 2002).

As we can see, children learn quite a lot about intention during their preschool years. They go from, at age 3, seeing intention as action, to at age 5, seeing it as a mental representation that differs from action, desire and outcome.

**Preschoolers' understanding of intentional versus unintentional behavior**

We have just seen what preschoolers know about intention and its relation to desires, action and outcome. What do they know about the difference between intentional and unintentional behavior? I review studies on this topic next.

In one study (Shultz, Wells, & Sarda, 1980), the authors had 3-, 5-, and 7-year-olds perform intentional actions and mistakes (mistakes were elicited by having children perform tasks at which they were likely to fail because they were difficult). The authors asked whether the children meant to do the mistakes, such as say a tongue twister incorrectly. Children of all 3 age groups said their actions were intentional more for the intentional actions than for the mistakes.

In a second study, Schultz et al. (1980) elicited movements in children that were either intentional or unintentional (e.g., a knee jerk brought on by tapping below the knee versus an intentional movement of the leg) and again asked, "Did you mean to x"? The 5- and 7- year-olds (but not the 3-year-olds) said "yes" more to the intentional than to the unintentional behaviors.

Why did 3-year-olds perform differently in the two studies? They seem to understand the difference between mistakes and intentional actions but not between passive movements and intentional actions. It could be that because the outcome in the first study (messing up on the tongue twisters) obviously conflicts with the original desire (saying them correctly), the "matching rule" strategy predicts the correct answer, whereas in the second study there is no relevant desire and the outcome is emotionally neutral. Alternatively, perhaps the 3-year-olds were simply less interested in the second study, as it involved such mundane tasks as moving a leg, as opposed to Study 1, which involved more exciting tasks like saying tongue-twisters and patting one's head while rubbing one's belly. Finally, it is possible that mistakes, which are the unfulfillment of a definite intention, are easier to differentiate from intended acts than are passive movements, which are performed without any intention. Perhaps there is something about trying to do something and failing that is more attention-getting and easier to grasp than doing something without having an intention at all.

In a somewhat similar study, Schultz and Wells (1985) also endeavored to test children's ability to discriminate between intentional and unintentional actions. They played a game with 3-, 7-, and 11-year-olds in which actors shot a ray-gun at different-colored targets. Colored cards indicated which target the actor was to shoot at and a light on the target indicated which target was actually hit. Children observed the card being chosen, the actor shooting the gun, and a particular target being hit (the outcome). Sometimes the intention matched the outcome and sometimes it did not match. Children were asked whether the actor hit the one he hit on purpose. Even the 3-year-olds were able to determine when the outcome was brought about intentionally by matching the intention to the outcome (when they matched, the behavior was seen as intentional, but when they did not match, it was not). Thus they showed an understanding of the

motivational aspect of intention (that the intention to act is usually directed toward a particular goal).

In another study addressing preschoolers' understanding of the intentionality versus non-intentionality of behavior, Smith (1978) showed 4-, 5-, and 6-year-old children various behaviors on videotape. Some of the behaviors were intentional (such as doing arm exercises and chewing something) and others were unintentional (such as sneezing and saying "ow" in response to a poke in the ribs). Participants were asked whether the actor was "trying to" and whether she "wanted to" do the actions. Four-year-olds tended to say that all the actions were intentional, whereas the older children said that only the voluntary behaviors were intentional.

However, the reason 4-year-olds performed worse on Smith's (1984) study than in the ones previously described may be that something about the cues to intentionality on the videotapes was amiss; several of the involuntary behaviors were judged to be voluntary by many of the adults. Also, since the videotapes contained several behaviors other than the one in question, the younger children may have misunderstood which action the experimenter was referring to. For example, in the sequence in which the woman was poked in the ribs and said "ow," these children may have thought the experimenter was asking about whether the person poking the woman in the ribs meant to do it, and not whether the woman meant to say "ow."

Most of these studies show that in certain simple cases 3-year-olds are able to detect intentionality of action using the "matching rule." Older children can make use of the rule in more complex situations. Although this does not require an understanding of intention as different from desire, nor does it include the important component of belief, it does require an understanding of the motivational component of intention. None of the studies reviewed in this section were designed to tease apart children's understanding of

desire from that of intention or to require consideration of the actors' beliefs; therefore we cannot know at this point just how mature preschoolers' conceptions of intentional versus non-intentional acts are. In Smith's study (1978) 5-year-olds showed a more mature concept than 4-year-olds in that they, but not younger children, knew which behaviors in the video were intentional, but the experiment's flaws seriously undermine the conclusion that 4-year-olds tend to interpret all actions as being intentional.

Overall, then, we have seen that Malle's (1999) model of how we explain human behavior stands up to conceptual inquiry as well as to empirical inquiry with adults. In addition, I have shown that preschoolers have some facility with various aspects of the model, such as a nascent understanding of intention and related mental states (e.g., desire, etc.) and the ability to choose and provide causal explanations for various actions and states of the world.

The overall question in the present research, one that previously has not been addressed effectively, concerns children's ability to shift their mode of explanation according to the intentionality of the behavior they seek to explain. Does this ability appear during the preschool years? If so, when does it appear, and how should its development be characterized?

## OVERVIEW OF THE PRESENT RESEARCH

To address this question empirically, I conducted two studies that were similar to those by Malle (1999) and White (1991) in the overall question as well as the basic design. To review, in Malle's Experiment 1 (1999), as well as in White's study (1991), adults were presented with causally ambiguous actions explained by reasons or causes. Based on the type of explanation given, they rated the intentionality of each behavior. Relations between intentionality judgments and the given explanation type were assessed.

Similarly, in my own Experiment 1, I assessed children's judgments of the intentionality of actions given cause or reason explanations. Young children are concrete thinkers in many ways, however, and thus find it easier to express judgments by pointing to one out of two illustrated scenarios than by verbalizing one of two abstract alternatives (intentional versus unintentional) that cannot effectively be represented through drawings (Browne & Woolley, 2004). Therefore, instead of presenting explained behaviors one at a time, I presented them two at a time, one with a reason explanation attached and the other with a cause explanation. Children were asked to indicate which of the two actions was done on purpose[3].

Recall that Malle's second study (1999) addressed the same basic concept as did his first study but differed in terms of which of the two pieces of information (intentionality of the action, type of explanation) was given by the experimenter and

[3] The reason I chose to ask children to point out the action that was performed on purpose in every trial, as opposed to asking them which action was not performed on purpose in half the trials, is that in my experience young children tend to get confused when the test question is not consistent for each trial but the possible responses are the same. Apparently it is not easy for children to "switch gears" in order to reason about opposite concepts between questions.

which was requested from the participant. Likewise, in Experiment 2 here children were presented with the intentionality of actions and were asked for explanations.

However, whereas Malle (1999) presented obviously intentional or unintentional actions (e.g., grinding one's teeth during sleep or asking a friend to come over for lunch) and assessed whether participants' explanations were reasons or causes, I presented the same causally ambiguous stories as in Experiment 1 and told participants whether they were done on purpose or not. I then asked them to choose between a reason and a cause explanation by pointing to the appropriate illustration. This design allowed me to compare performance in the two studies directly, as well as to ensure that children's explanation choices were based on intentionality as explicitly provided by the experimenter. If children's explanation choices were derived from their own judgments of the intentionality of particular actions, there would be more than a remote possibility that between-subject variability in the main dependent variable (which explanation was chosen) would include variability in children's tacit judgments of how purposeful the given actions are. Directly telling children whether or not each action is intentional therefore would presumably result in "cleaner" data.

With respect to performance in Experiments 1 and 2 considered together, it was difficult to come up with precise expectations regarding the age at which children would show an understanding of the relation between intentional action and mental explanation and between unintentional action and physical explanation. As a whole this research suggests that 3-year-olds have some understanding of different types of causality but are "bad explainers" and depend on outcome information to make judgments about intention. As the research conducted here gave no a priori desire information and the outcomes were mostly neutral, it seemed quite unlikely that this age group would succeed. Therefore, 3-year-olds were not included as participants in the present experiments.

Four-year-olds have a more mature concept of intention and explanation; however, they apparently do not understand that two people with different intentions can perform the same action, which is an important aspect of the present tasks. What little research included 5-year-olds indicates that, crucial to this research, a single action can be motivated by different intentions. Nevertheless, the tasks here required systematic coordination of intentionality of action and type of explanation, thus it seemed that this ability might continue to develop beyond age five. In addition to 4- and 5-year-olds, 6-year-olds were thus included as participants.

## EXPERIMENT 1

### Method

#### *Participants*

Seventy-six children (approximately half girls) participated. Twenty-nine were 4-year-olds (*M* = 4;5, range = 4;0-4;11), twenty-five were 5-year-olds (*M* = 4;6, range = 5;0-5;11), and twenty-nine were 6-year-olds (*M* = 6;6, range = 6;0-7;0). Names were obtained from birth records kept on file at a university research laboratory and children were recruited through letter and telephone contact with the parents. Most participants were White and middle class, but several ethnic groups were represented, including Latino, African American, East Asian, and South Asian.

#### *Stimuli and Design*

Each participant was presented with six illustrated stories featuring protagonists performing actions. Two versions of each story were presented to children in each of the six trials: one in which the action was explained by giving a reason, and one in which the (same) action was explained by giving a cause. The versions differed only in the name of the protagonist performing the action and the explanation for the action; the actions themselves, the illustrations, and the gender of the protagonists were the same.

It was important that the actions in the stories be causally ambiguous. That is, the action itself, without explanation, should be perceived as being approximately equally likely to be performed intentionally as unintentionally. Causally ambiguous actions were chosen by presenting eleven adults with a list of 10 actions that seemed to be causally ambiguous and having them rate the intentionality of the actions themselves, without explanation, on a scale of -5, “definitely unintentional” to +5, “definitely intentional.” A

zero rating represented causal ambiguity. The six actions whose average of the eleven ratings was equal to 0, +1, or -1 were used.

In order to control for possible effects of story order and potential biases in children toward the intentionality or non-intentionality of particular actions[4], two variables were counterbalanced or randomized. First, story order was counterbalanced between participants. Within participants, the order in which the two types of explanation for actions (reason versus cause) were presented within the answer choices alternated.

An issue that deserves special consideration involves the particular words used to represent the concepts of intentionality and non-intentionality in both Experiments 1 and 2. Obviously this is an important matter, as the validity of the data generated by this research rests on conveying to children the meaning of the words as well as the relation between them.

To convey intentionality of action I chose to use the phrase "on purpose" for a few reasons. First, in response to an informal, free-response question directed to the parents before the commencement of the experimental session, they were most likely to include the term "on purpose" in reporting which words/terms their child uses to indicate the intentionality of an action compared to other terms. The second most reported phrase was "meant to," which is used in part of the test question, as explained later. Although some other studies have used the term "try to" to indicate intentionality (e.g., Astington, 1991), this term is problematic in addition to rarely being uttered by children (according to their parents). The main issue involves the term's implication concerning whether the

[4] Although adults had previously judged the stories to be completely or nearly causally ambiguous, it seemed possible that children could differ slightly in their perceptions of whether certain actions are more likely to be either intentional or non-intentional.

action was successfully executed. According to common usage, "try to [action]" implies a failed effort. For example, one might say, "I tried to kill the mosquito." This statement implies that the speaker desired, attempted, and endeavored to catch the mosquito, but her efforts were somehow thwarted. "On purpose," on the other hand, clearly communicates that the desired action was successfully performed: "I killed the mosquito on purpose." Because the actions in the study are successfully performed, using "on purpose" in this study is arguably a better alternative than "tried to."

To indicate the non-intentionality of actions in the experiment, I used the term "not on purpose". Although the most commonly used word/term used by children to indicate the non-intentionality of actions (as reported by their parents) was "by accident," this term, with its connotation of a negative outcome, does not quite represent the concept this research is meant to tap. The implication of a negative outcome when "by accident" is used (especially by children) was confirmed by parents, who nearly all indicated that their children use "accident" primarily in the context of fighting with siblings or doing something for which they could get into trouble (e.g., hitting, knocking over sand castles, etc.).

A final reason to use "not on purpose" is to clarify that intentional and unintentional actions are mutually exclusive categories whose meanings are precisely opposite. When a phrase is contrasted with its negation, with the two members of the pair explicitly referring to each other, the pair more clearly communicates an "either-or" construct – in this study represented by the dichotomous forced-choice nature of the questions – than does a pair whose members, though antonymous in some ways, do not explicitly refer to or negate each other. To cite another example, the words "tied" and "untied," referring to shoelaces, are clearly mutually exclusive and opposite in meaning; every pair of shoelaces in the world is either tied or untied (ignoring in-between cases

such as laces that are extremely badly tied). However, even though, for example, "tied" and "loose" can be construed as antonyms because they are contrastive when referring to shoe laces, "loose" may project connotations (perhaps "messy" or "unkempt") that "untied" does not have.

***Procedure***

Children were tested individually in a university laboratory in a mid-sized Southwestern city. First, children underwent a warm-up session consisting of a clarification of the terms/concepts "on purpose" and "not on purpose." The first five minutes or so were informal; the experimenter tailored her own portion of the interaction to the children's responses, thus the structure and content of this part of the session varied somewhat from child to child.

First the experimenter simply talked with the child about the meaning of "on purpose" or "not on purpose." As stated earlier, almost every child had some notion of the "on purpose/not on purpose" dichotomy, but these notions were often wrong or incomplete. For example, in response to the question "Do you know what it means to do something that's "on purpose/not on purpose?" many children at first explained the concepts with reference to a typical domestic scene in which sibling rivalry, blame, punishment, exoneration, and apology are interrelated. For example, with reference to "on purpose" several children said something like "when you do something mean and don't say sorry." One girl said "I hit my brother and he got mad and he told on me, and I got in trouble and I had to say sorry." For "not on purpose" many children's responses involved avoiding punishment.

When children expressed either an incorrect concept or no concept of the "on purpose / not on purpose" distinction, the experimenter explicitly corrected children's concept or explained what "on purpose" and "not on purpose" means. Again, the

experimenter's part of the conversation varied according to the child's present understanding and the situation as a whole, but emphasis was placed on "doing something that you were *trying* to do, something that you thought you were going to do" versus "something that just happened, that you weren't trying to do." Together the experimenter and the child thought of a few valid examples of each. If the child did not come up with any examples at the experimenter's request, the experimenter described several actions, such as stepping on someone's toe, playing, and breathing, and asked whether they were on purpose or not. For these questions feedback and explanations were given. The goal was to train children on the relevant concepts as well as possible before beginning the warm-up questions.

After this informal discussion, the experimenter asked four standard warm-up questions about whether certain actions are on purpose or not on purpose: 1) the experimenter walking around the table, 2) the experimenter bumping into a chair, 3) dancing[5], and 4) sneezing. These questions were meant to be fairly difficult. I wanted to emphasize to children that sometimes it took considerable effort to decide whether an action is on purpose, as they tended to begin the session by acting as if the decision were easy to make, even if they didn't really understand the concept. In addition, it made sense to ask them about actions that (except for one) had neutral outcomes, which was contrary

[5] Interestingly, in response to the question, "Is dancing on purpose or not on purpose?" two children said that, although it is usually on purpose, if someone were "walking really gracefully" or "going like this" (moving her arms around) she would be dancing, even though it would not be on purpose and, presumably, not known by the actor. This concept of "dancing" suggests that there is a fact of the matter regarding whether or not a certain type of movement counts as "dancing," rather than that the concept "dancing" involves the actor's intentions as well as others' perceptions. In other words, in contrast with adults' concept of dancing as an artifact category, these children treated "dancing" as a natural kind.

to their common view that "purpose" and "accidental" are intrinsically related to a negative outcome. Children received feedback and explanation after each question.

The two versions of each story were then read to children in succession, presenting different explanations (cause or reason) for each version. (See the actual stories used in the Appendix.) In the *blocks* story, for example, while showing the child one of the action pictures the experimenter said, "I'm going to tell you two stories. This one is about Cathy. She was playing with some blocks and she made a tower. Then, she knocked over the tower. She knocked over the tower because she tripped on it. [Last sentence was repeated.]" The other picture (identical to the first) was placed on the table next to the first one. "I'm going to tell you another story. This one is about Amy. She was playing with some blocks and she made a tower. Then, she knocked over the tower. She knocked over the tower because she wanted to make something else. [Last sentence was repeated]."

After the actions and explanations were presented, there was a "memory check"[6]: "Point to the one who knocked over the tower because she wanted to build something else." After the child pointed: "Now point to the one who knocked over the tower because she tripped on it." In the infrequent event that a child erred on a memory check (97% correct overall), the experimenter corrected him and repeated the question until the participant responded correctly. The dichotomous, forced-choice test question was then asked as the experimenter pointed to each illustration in succession: "Which one meant to knock over the tower? The one who knocked over the tower because she tripped on it, or the one who knocked over the tower because she wanted to build something else? Point

[6] The "memory check" questions were asked in order to make sure the children were paying attention to the story as well as to emphasize its most important aspects (the explanations in Experiment 1 and the intentionality of the action in Experiment 2), not primarily to test in a meaningful way children's memory for what they had just been told.

to the one who knocked over the tower on purpose."[6] No feedback was given for the test questions.

**Results and discussion**

***Responses to warm-up questions***

For these questions, responses of participants in Experiments 1 and 2 were combined for the purpose of analysis because the two studies included the same warm-up session and questions. Each participant was given a score, 0-4, representing the number of warm-up questions the child answered correctly. Overall, children's scores were fairly low ($M$ = 2.6, or 66%). As mentioned earlier, however, these questions were meant to be more difficult than the ones in the main experiment. It was necessary for them to counter the ubiquitous notion among participants that the question of intentionality is always about negative outcomes and blame. According to parents, all except one of the participants used the on purpose/not on purpose dichotomy at home, and in every case the usage was said to be in the context of negative outcomes such as hurting someone or breaking something. In the test stories, however, none of the outcomes is inherently negative. Therefore, the warm-up questions served as training trials in that three of them involved neutral outcomes and children received corrective feedback and explanation following responses. Thus, by the time they reached the test trials many children's understanding of the dichotomy had improved.

Looking at the warm-up scores per age, the mean for four-year-olds was 2.3, or 58%, for five-year-olds it was 2.6 or 66%, and for six-year-olds the mean was 3.18, or 80%. The fact that the ages differed in their understanding of intentionality was considered to constitute part of the variability in the major question addressed here. That

is, the major question is not based on the premise that the participants have equally correct or complete concepts of intentionality before they begin the test trials.

In Experiment 1, children with warm-up scores of 2 or higher were included in the rest of the analysis. Although a score of 2 indicates chance level performance, if cases with scores of 2 were excluded there would not be data from enough participants to analyze it with much power or validity. Also, excluding those with scores of 0 and 1 ensured that the analyses did not include data from children who had reversed the concepts (on purpose and not-on-purpose) in their minds. After cases with the lowest scores were excluded there remained 21 four-year-olds, 21 five-year-olds, and 20 six-year-olds.

***Stories***

One story, *bug*, was eliminated from further analysis due to an experimenter error in which the story was presented to children as unintentional in three of the four "orders" and as intentional in only one. Responses on the test question for the other five stories were submitted to multiple correlations. These responses in most of the stories were significantly positively correlated with those of every other story, and the two correlations that were not statistically significant were also positive, though small. Due to the overall positive relations of responses across the five stories, I considered them as a group for the rest of the analyses.

***Test question (purpose judgments)***

After both versions of each story were read to the child, he was asked which character [performed the action] on purpose. Each participant was given a score of 0-5 indicating the number of test questions answered correctly. A *t*-test revealed that the overall mean (3.6, or 72% of the maximum score) was statistically greater than chance, *t*

(61) = 5.57, $p < .001$, but not at ceiling level, suggesting that the task was appropriate for this age range (4-6).

Next I addressed the question of age as a factor in performance. A one-way ANOVA showed that the age groups' means (2.8 or 56% for 4-year-olds, 3.6 or 72% for 5-year-olds, and 4.5 or 90% for 6-year-olds) differed significantly, with performance improving with age, $F$ (2, 61) = 7.9, $p < .01$. See Figure 1 for a depiction of these means and standard errors (and the corresponding means and SE for experiment 2). To analyze this age effect further, I performed post-hoc multiple comparisons (Tukey HSD) contrasting the performance of each age with that of the others. Although neither the 4-year-olds' nor the 5-year-olds' performance differed significantly from that of the next oldest group (4 versus 5, 5 versus 6), 6-year-olds performed significantly better than did the 4-year-olds, $M$ difference = 1.7, $p < .01$.

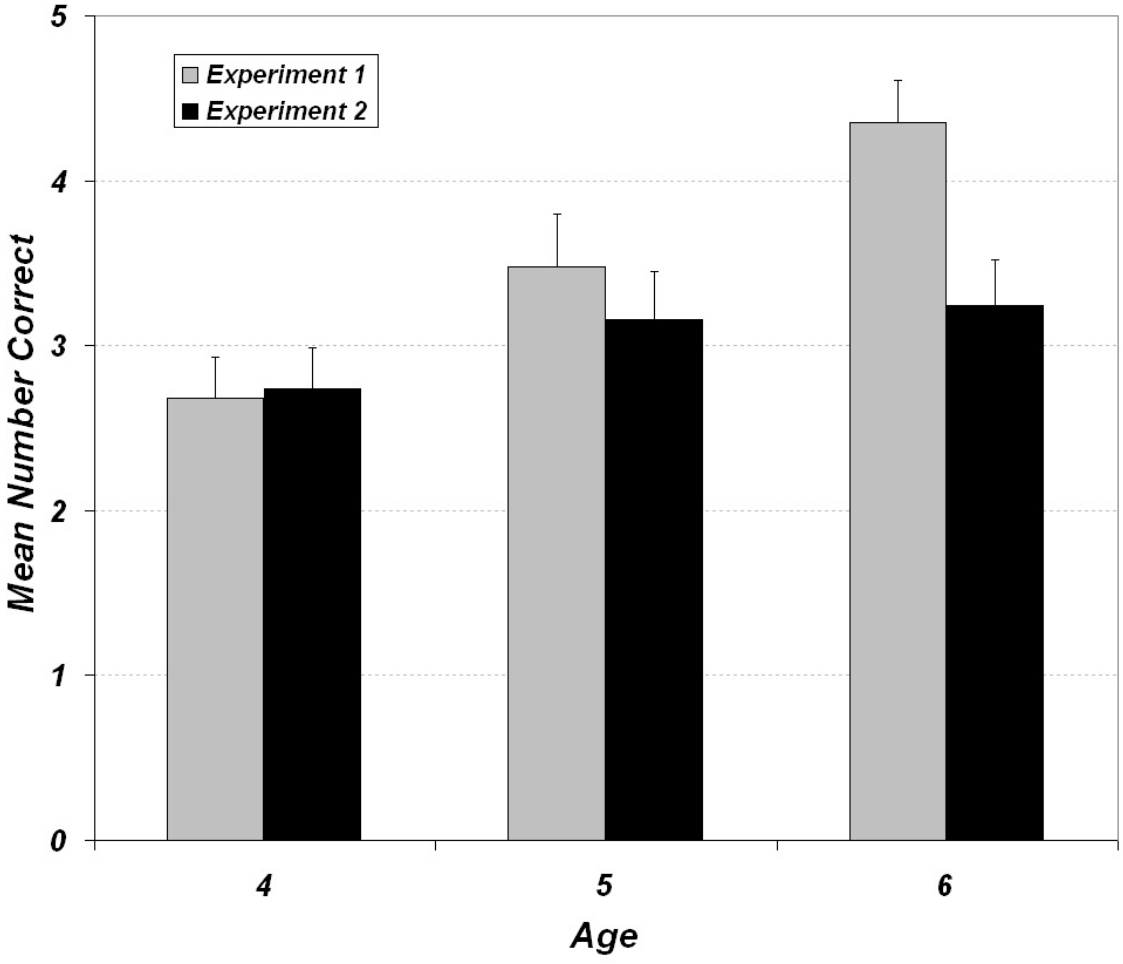


Figure 1: Test question performance by age and experiment.

Despite the fact that a direct comparison of 4-year-olds' and 5-year-olds' scores indicated that they were not significantly different, 4-year-olds' performance was no better than chance level, whereas5- and 6-year-olds' performance exceeded chance by a fairly wide margin (for 5-year-olds t (20) = 2.9, $p < .01$; for 6-year-olds $t$ (19) = 8.5, $p < .001$). Thus, it seems that near the beginning of their sixth year children acquire some understanding of the relation between the intentionality of action and the kind of

causality (mental versus physical) that should be applied to the explanation type, and that this understanding continues to develop to nearly ceiling level during the next two years.

## EXPERIMENT 2

In this study the experimenter read to children six stories about a character performing an action, told them whether or not the action was on purpose, and asked them to choose between a cause explanation and a reason explanation for the action.

### Method

#### *Participants*

Seventy-five children (approximately half girls) participated. Twenty-four were 4-year-olds (*M* = 4;6, range = 3;11-4;11), twenty-four were 5-year-olds (*M* = 5;6, range = 5;0-5;11), and twenty-seven were 6-year-olds (*M* = 6;6, range = 6;0-7;0). Names were obtained in the same manner as in Experiment 1 and the ethnic makeup of the participants was approximately the same.

#### *Stimuli and design*

As in Experiment 1, each participant was presented with six illustrated stories depicting protagonists performing causally ambiguous actions (the same stories and drawings as in Experiment 1). In Experiment 2, however, there was only one version of each story instead of two. After each story was read to the child and he was told whether the action was on purpose or not, he was asked to choose between two drawings representing possible explanations for the action.

To control for order effects and potential intentionality biases, again a few variables were counterbalanced. First, the order of the stories was counterbalanced across participants. Then, the intentionality of each action presented (whether the action was intentional or not) was counterbalanced between participants so that the action in each story was said to be intentional for half of the children and unintentional for the other

half. Within subjects, the intentionality of the first answer choice for the test question ("Point to the one that shows why she did it") was counterbalanced.

It was important that the drawings representing the two explanations in each trial not "give away" the intentionality of the action they explained. If children tended to perceive either of the two explanation drawings in a trial as somehow depicting more intentionality than the other, they could respond to the test question ("Point to the one that shows why she did it…") on that basis instead of by considering the explanations themselves and whether the action was said to be intentional or unintentional. The goal was for the illustrations to represent the explanations so they could serve as a concrete icons for the children to point to after hearing what they represent, but not to indicate the actions' intentionality directly.

To address this concern I showed ten 4- through 6-year-olds the pairs of explanation pictures used for each story. After I told them what the action in the pictures was I asked which depicted child did it on purpose. For example, for the *jungle gym* story, in which a boy drops down from a monkey bar, the picture for the cause explanation shows a boy hanging from a bar with one hand, with curved lines over the hand showing movement. The other picture shows the same thing, but with a swing within the boy's sight. The first picture is supposed to represent the cause (the boy's hand slipping off the bar) and the other is supposed to represent the reason (the boy dropping down from the bar in order to play on the swing). The question was: "These two boys dropped down from the bar. Which one did it on purpose?" For each pair of pictures approximately half of the responses were correct. Thus, children could not infer the intentionality of action based on the explanation pictures.

***Procedure***

The warm-up session was the same as in Experiment 1. Following the warm-up, children received the six experimental trials. For each of the trials, the experimenter read a story about a protagonist performing an action. For example, "This is Frank. Frank was at a birthday party and someone gave him a balloon. Then he popped the balloon." Children were told whether the action was on purpose or not: "He meant to pop the balloon. He did it on purpose" or "He didn't mean to pop the balloon. He didn't do it on purpose." The "memory check" was next: "Did he do it on purpose, or not on purpose?" Feedback was given. For the two children who answered a memory check incorrectly the question was repeated until they gave the correct answer.

Then the test question was asked: "Why do you think he popped the balloon? Do you think he popped it because he wanted to hear the popping sound it made (pointing to the picture), or do you think he popped it because he dropped it on something sharp (pointing to other picture)? Point to the one that shows *why* he popped the balloon." Finally, children were reminded of the intentionality of the action: "Remember, he did it on purpose." For the test question no feedback was given.

The hypotheses for Experiment 1 applied to this study as well, but were additionally informed by the research about preschoolers' explanations. Recall that preschoolers are able to choose reasonable explanations out of several choices. However, choosing explanations in this experiment seems more difficult than that in the studies mentioned above. For example, in Springer and Keil's (1991) study many of preschoolers' explanation options were obviously wrong, such as explaining the color of flowers as *a little man who paints flo wers pink*. As in Experiment 1, it seemed likely that the development of choosing explanations based on the intentionality of action would not be complete at age five and would continue to improve throughout the child's sixth year.

## Results and discussion

### *Warm-up Questions*

As stated earlier, for these questions the data from the two experiments were combined. See the results for Experiment 1 for these data. Participants with warm-up scores of less than 2 were again excluded from further analyses. The remaining participants in Experiment 2 were 16 4-year-olds, 18 5-year-olds, and 24 6-year-olds.

### *Stories*

Because the *bug* story was excluded from analysis in Experiment 1, it was excluded here as well. Fortuitously, multiple correlations on the six stories indicated that *bug* was the only story that was negatively correlated with the others. Most of the other correlations (except 3 positive non-redundant correlations), though positive, were non-significant. To further investigate the relations among the stories I submitted them to a factor analysis. The criterion for retention of components was whether their eigenvalues were equal to or greater than 1.0. Two factors were extracted: *bug* was one factor, and the other stories comprised the second factor. Thus, data from the other five stories were included in the rest of the analysis.

### *Test Question*

Each participant was given a score of 0-5 indicating the number of test questions answered correctly. A *t*-test revealed that, as in Experiment 1, the overall mean (3.2, or 64% of the maximum score) was statistically greater than chance, $t\ (57) = 3.6$, $p < .01$, but not at ceiling level, indicating that this task was also appropriate for this age range (4-6).

The means for each age were as follows: For 4-year-olds $M = 2.8$, or 56%, and for both 5- and 6-year-olds $M = 3.3$, or 66%. Regarding age as a factor in Experiment 2, a

one-way ANOVA showed that overall age did not affect performance, $F$ (2, 55) = .66, $p$ > .05. Although there was no significant age effect, I performed post-hoc multiple comparisons (Tukey HSD) contrasting the performance of each age with that of the others to see whether the 4-year-olds' performance, which appears to be quite a bit lower than that of 5- and 6-year-olds, differed from them when compared directly. The comparisons comfirmed the ANOVA, showing that no age group performed significantly better than any other.

However, as in Experiment 1, the performance of these 4-year-olds were worse than that of the older groups in the sense that the youngest group's mean, 2.8, does not exceed chance level (2.5) whereas the older group's means do. For 5-year-olds $M$ = 3.3; $t$ (17) = 2.2, $p$ < .05, and for 6-year-olds $M$ = 3.3; $t$ (23) = 2.7, $p$ < .05. The $t$-values differ despite apparently equivalent means because the groups have unequal numbers of participants.

***"On purpose" versus "not on purpose" trials***

Because Study 2's design involved within-subject differences, with half of the trials featuring intentional actions and half featuring unintentional actions, it was possible to assess whether children found one type of action to be easier to reason about than the other, as well as to see whether responses on the two trial types were correlated.

To assess the effects of trial type (whether the action presented is intentional or not) and age on performance, as well as whether the effect of trial type varies by age, I conducted a mixed design ANOVA. This between-within analysis showed that there was no interaction between age and the two types of trial, $F$ (2, 56) = .11, $p$ > 05, indicating that the effect of trial type was similar at each age group. Specifically, scores on the *on purpose* trials ($M$ = 1.7, or 55%) were significantly higher than those on the *not on purpose* trials ($M$ = 1.4, or 45%). See Figure 2 for mean scores and standard errors. This

finding, in combination with the absence of an interaction between Age and Trial Type, suggests that even as children progress from responding at chance level to responding significantly above chance, reasoning about unintentional actions (at least in terms of explaining them) lags behind reasoning about purposeful actions.

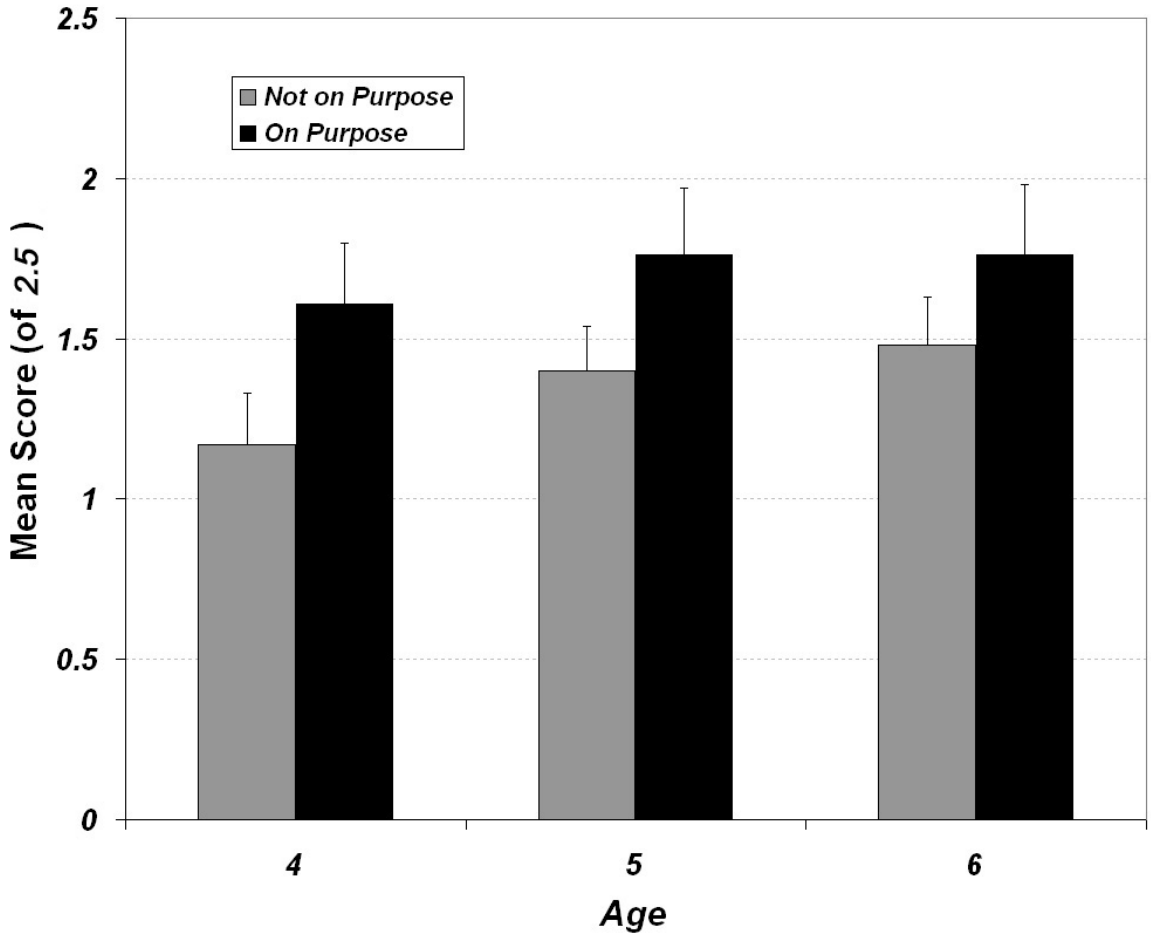


Figure 2: Performance on intentional and unintentional trials by age.

***Correlation between trial types***

To find out whether performance on one trial type (e.g., intentional action) had any bearing on that performance on the other trial type (unintentional action) I calculated the Pearson product moment correlation between them and found it to be quite small, $r$ =

.19, *ns*. The fact that responses on the *on purpose* and *not on purpose* items were uncorrelated was somewhat surprising, as they were meant to address different aspects of the same concept.

**Comparison of Experiments 1 and 2**

A two-way (Age X Experiment) ANOVA revealed that there was no significant interaction between experiment and age, $F(143) = 2.3$, $p > .05$. This finding suggests that performance in Experiment 2, which seemed to improve somewhat with age but showed no significant age effect when considered separately, had enough of a developmental trend to prevent an experiment X age interaction. Thus, the question of whether performance on Experiment 2 improved with age is not straightforward. When analyzed alone it does not show an age effect. However, as Experiment 1 shows a significant age effect and there is no interaction between age and experiment, it would not be accurate to consider performance on Study 2 not to have been affected by age.

## GENERAL DISCUSSION

### Summary of results

In this research I investigated young children's knowledge of the relationship between the intentionality of an action and whether it was generated by the actor's mental states or by a direct mechanical force. Four-year-olds seemed to possess little understanding of this relation, whereas 5- and 6-year-olds in both studies showed a moderate degree of understanding. Performance in Experiment 1 improved significantly with age, and evidence that performance in Experiment 2 improved with age was mixed.

In Experiment 2, the trials in which children chose an explanation for an intentional action were compared with those in which they chose an explanation for an unintentional action. Overall, children performed better on the intentional trials, and there was no significant interaction between age and trial type. These findings suggest that the difference between scores on the two trial types was similar across ages. Finally, performance on trials with intentional actions was uncorrelated with performance on trials with unintentional actions.

How do these findings relate to those of the most relevant studies reviewed earlier in terms of the age at which the knowledge assessed by my research is gained? Does past research serve in any way serve to explain or corroborate these results?

### The chronology of learning to explain behavior

In most of the relevant past studies participants were younger than those tested in the present research (e.g., Barsch & Wellman, 1989; Hood & Bloom, 1979; Kalish, 1998), and therefore the findings do not directly speak to the present issue as a whole. The authors of these studies were mainly concerned with young children's understanding of some of the important basic components of the relatively complex ability addressed

here. Some of these components include causal explanation in general, the understanding of intention as a mental state, the distinction between intentional and unintentional behavior, and the ability to be flexible in explaining various types of actions (i.e., by referring to intentional, biological, or physical forces). Thus, although there is enough past research to get an idea of the nature and timing of young children's abilities that are relevant to the present research, the extent to which the findings of past studies can shed light on the present findings is somewhat limited.

One question involves the failure of the 4-year-olds in the both of the present studies to show understanding of the relation between intentionality of action and type of explanation in their responses. Two studies (Kalish, 1998; Schult & Wellman, 1997) indicate that 4-year-olds are capable of explaining human behavior flexibly. That is, the studies show that these children understand that non-intentional forces can affect, or even generate, human action.

However, the present studies arguably require a more advanced, or perhaps more specific, understanding of the relationship between action and explanation than do these previous studies. First, as discussed earlier, Kalish's study (1998) does not really address the understanding of *causal* explanation. In addition, in that study all the actions involved an obvious intentional component (intending and attempting to perform the action to fulfill a specific desire), so the only aspect of each scenario that was *not* intentional was the outcome. Knowing simply that intentional actions are sometimes thwarted by non-physical forces seems easier than understanding that 1) causally ambiguous actions can be generated either by mental or by physical forces, and that 2) the type of force generating the action dictates whether or not the action was on purpose as well as what type of explanation should be given.

Similarly, although Schult and Wellman's studies (1997) also provide evidence that 4-year-olds are not constrained to psychological explanations for human behavior, again the tasks in their research seem easier than the present ones. In those studies the stories with intentional actions ended with an inherently positive outcome, whereas the stories with unintentional actions ended with an inherently negative outcome. Children could use desire words like "want" in the explanation whenever the actor obtained what she wanted (using something like the "matching rule"), and could avoid such words when the outcome did not match the actor's desire. This strategy would lead to explaining actions in different ways without necessarily understanding the underlying factors that generate intentional and unintentional behavior.

By contrast, the tasks in the present studies do not provide major clues to children about how to explain an action beyond the statement of whether or not it is on purpose, as the actions are causally ambiguous and the outcomes are the same whether caused by an intentional or unintentional action. Thus the poor performance of the 4-year-olds in this difficult study does not necessarily contradict the findings of relevant studies done previously.

Not much is known, however, about the development between ages 5 and 6 of intention- and explanation-related knowledge because most of these studies test relatively simple abilities. Thus 6-year-olds are usually not included within the participants. It seems that currently the most common approach to developmental research related to the present studies is to find the very earliest that children begin to demonstrate a modicum of understanding of a particular concept or ability without exploring how the concept or ability continues to develop. Perhaps investigation of children's concepts and abilities should be extended to older ages, as concepts and abilities are not usually discrete entities that are acquired all at once. On the other hand, perhaps other researchers have found

little relevant development in pilot tests comparing 5- and 6-year-olds and thus focus on younger children.

As discussed earlier, although there was no interaction between experiment and age, when considered separately Experiment 1 showed an age effect but Experiment 2 did not. It was expected that the understanding addressed here would improve with age, so the real question concerns why there was not more of an age effect in Experiment 2. Recall that in this experiment children were told whether actions were on purpose or not on purpose, and they chose one of two possible explanations for the actions, either a reason or a cause. Perhaps there is some limiting factor present in Experiment 2, but not in Experiment 1, that is unrelated to the main question in the research. One candidate for such a limiting factor might be the ability to "switch gears" from reasoning about intentional actions to reasoning about unintentional actions. The order of the two trial types was randomized, so children did not know before any given question whether the action would be on purpose or not. After learning the intentionality of the action, participants had to change their mode of reasoning quickly, from one type of action, causality and explanation to the other. Although this processing variable seems as if it would improve during the preschool years, perhaps the time course of this ability's development is on a larger scale, such that improvement can be detected over many years instead of from one year to the next.

Alternatively, perhaps the 5- and 6-year-olds were already at adult levels for this particular study. As I will discuss in the following section, even adults tend to have a natural bias toward "psychologizing" actions, assuming purposefulness and reason explanations when in doubt or when the issue is not being considered.

**Reasoning about intentional and unintentional actions**

Recall that in Experiment 2, in which children chose explanations for actions featured in the stories, there were two trial types differing by whether the experimenter announced that the action in the story was on purpose or not on purpose. Although both types of trial were designed to assess the general ability to match the intentionality of actions with appropriate explanation types, children performed better when they reasoned about the "on purpose" items than when they reasoned about the "not on purpose" items, even though across participants the same stories were used for both trial types. In other words, children were more likely to err by choosing reasons for unintentional actions than by choosing causes for intentional actions.

This finding suggests that the children were biased toward explaining behavior in terms of desires and / or beliefs, even when the experimenter clearly stated that the action was not performed on purpose. Whereas other studies had demonstrated that children are capable of explaining different types of behavior using different modes of causality, to my knowledge this is the first study directly comparing children's explanations for intentional actions with those for unintentional actions such that one can see whether mistakes in explanations were distributed evenly between the two types of action, or whether explanations overall tended toward either reasons or causes.

Is this finding consistent with the results of the other research reviewed here? Indeed, a few of the studies discussed earlier are suggestive of such an intentional bias. First, McCabe and Peterson (2001), as well as Hood and Bloom (1979), found that in natural speech 2- through 6-year-olds generally talk more about mental causality than about physical causality, indicating that young children gain more experience with the latter type of causality than the former, and perhaps also that young children tend to focus on the social world, with its characteristic mode of causality, more than on the physical

world. In addition, in Smith's study (1984) 4-year-olds looking at both intentional and unintentional videotaped actions consistently responded that the actions were on purpose, seeming to reveal a default strategy of assuming intentionality when one is in doubt or when one does not give the decision much consideration.

In fact, even Malle's (1999) adult participants seemed to be biased toward judging actions to be intentional. In Study 2, when they judged the intentionality of actions based on given explanations, they did in fact conform to Malle's general model of behavior in that their average intentionality ratings for actions explained with reasons was higher than the average intentionality ratings for actions explained with causes. However, whereas the mean of their intentionality ratings for actions explained with causes was 8.1, which was close to maximally intentional, the mean of their intentionality ratings for actions explained with causes was 4.1, which is relatively closer to causal ambiguity. In other words, even though the actions were meant to be causally ambiguous, and each was explained with a cause for half the participants and with a reason for the other half, adults tended to avoid judging actions to be clearly unintentional. It is not possible to statistically evaluate this claim, as Malle neither analyzed nor discussed the issue.

Another, less interesting possibility regarding the discrepancy between intentional and unintentional trials in Experiment 2 of the present research is that children find it more difficult to reason about actions described in terms of a negation (i.e., *"didn't* mean to," "*didn't* do it on purpose") than actions described in simple, positive terms (i.e., "meant to," "did it on purpose)." However, in a series of studies in which children were required to make dichotomous decisions about whether actions were "magic" or "not magic," Woolley and I found that even children who had just turned three were as likely to say that the items were *not* magic as they were to say that they were magic (Browne & Woolley, 2004). Also, the fact that the discrepancy between trial types in the present

research corroborates other research findings in which people seem to resist construing actions as entirely unintentional (e.g., Hood & Bloom, 1979; Malle, 1999; McCabe & Peterson, 2001; Smith, 1984) argues for the interpretation that people's construals of the intentionality of actions are made against a tacit background assumption that most actions are on purpose or generated by mental states.

This "bias" toward seeing actions as intentional or as able to be explained using reasons is not altogether surprising, as everyday experience suggests that we as people are more focused on and interested in actions that reflect something about the mental states of the actors than actions that reflect non-intentional forces. In Smith's words (1984) we tend to "cognize the behavior stream" in terms of a series of intentional actions (p. 736). It seems that in order to do this one needs to ignore physical details of how the actions are carried out (something Fune could not do, Borges, 1967), many of those insignificant-seeming details being unintended movements or actions.

A possible factor contributing to this intentional bias is that we tend to think of our "selves," or our identities, as being something other than our bodies. Bloom (2004), for example, following DesCartes, argued that we tend to identify with ourselves primarily as incorporeal beings who are *contained* in our bodies or who *possess* our bodies, indicating that on some level, regardless of our religious beliefs, we think there is a self (or soul, or mind) that is at least theoretically separable, or even separate, from our physical manifestations.

This attitude is evident when we say "use your brain," as if the brain were something that we use as a tool, like a calculator, or when people debate about whether some mental illness is a "physical" disease, as if there could be any mental state, mental dysfunction, or mental event that was not inextricably tied to corresponding "events" in the physical brain. This intimate identification with the mental, or psychological, self as

an autonomous being apart from the body might lead us to discount or fail to acknowledge some of the unintentional actions performed by our bodies (not by our "selves"!) without mediation of the will.

In other words, perhaps our intuitive dualism leads us to feel more comfortable explaining our own and other people's behavior in terms of our mental states (which are closely identified with the "self") than in terms of mechanical forces (that in most cases directly affect the body, which is less identified with the self). Although people do not seem to have much practical trouble explaining unintentional actions with reference to physical forces, we might feel *slightly* uncomfortable viewing ourselves or other people as conforming to the same laws as other physical objects such as billiard balls and planets. This special focus on intentional behavior could lead people to deal with ambiguous cases by erring in the direction of intentionality and purpose, thereby coming up with "false positives."

It should be noted that this bias does not contradict Malle's claim that we tend to view intentional actions as being caused by the actor's mental states and unintentional actions as being caused by direct, non-mental forces. The claim is simply that intentional actions gain more attention from us than unintentional ones, and that there is a slight bias to interpret actions intentionally. After all, the model is supposed to describe a cognitive, behavioral, and perhaps emotional tendency, not an invariable or inevitable pattern.

In light of this discussion of some of the cognitive-emotional factors involved in explaining behavior, perhaps Carey and Spelke's (1994) "entity view" of how we decide what type of causality to refer to in our explanations is not entirely incorrect. A hypothetical strong version of the view, claiming that type of causality is necessarily linked to the entity being explained (e.g., mental causality for people, physical causality for computers) has been disproved by several studies showing children's and adults'

flexibility in explaining human action (e.g., the present research; Kalish (1998); Malle, 1999; Schult & Wellman, 1997). Again, however, the present findings (that children seem to err on the side of "over-psychologizing" human behavior), along with similar findings from other studies such as those previously mentioned and our intuitive views about our selves and our bodies, suggest that we do have a natural tendency to explain human behavior in terms of mental states. The extent of the naturalness and flexibility of attributing the movement or "behavior" of inanimate objects to mechanical causality is beyond the scope of this research, but I would not be surprised if it turned out that both children and adults tended to explain the behavior of "intelligent" objects such as computers in terms of its desires and even its beliefs (e.g., "It's trying to open the program but it thinks I'm still using the other one.")

**The nature of the ability to explain and interpret human behavior**

In Experiment 2, not only was performance on the intentional trials better than that on the unintentional trials, but performance on the two trial types did not correlate. Together these findings lead to the question of whether the ability to explain actions based on their intentionality consists of a unified set of rules or relations or of two different "modes" of reasoning or knowledge appropriated for the task.

Traditional, domain-general views of cognitive development, adult cognition, and attribution theory would maintain that the ability to explain human behavior gradually improves through social experience and the development of general reasoning skills. However, more recent theory within psychology would suggest a different view.

In the last fifteen years or so, the idea that the nature and development of our knowledge, thought processes and development are at least partially domain specific has garnered increased attention. According to the domain-specific view of cognition and development, learning takes place within several distinct systems of knowledge, each

with its own ontology and causal laws. Two of the most commonly cited candidates for domains are basic knowledge about people's mental states and behavior ("intuitive psychology") and basic knowledge about physical objects ("intuitive physics"). (For a thorough discussion of the domain specificity view, see the chapters in Hirschfeld & Gelman, 1994.)

Consistent with the domain specificity view of cognition and cognitive development is the proposal that the task of explaining human behavior, although it is a crucially important, ubiquitous, and almost effortless skill, is best conceived as consisting of two separate modes of reasoning. Under this view, we decide whether or to what extent a behavior is intentional and come up with an explanation for it by consulting and coordinating two of our main knowledge domains, intuitive psychology and intuitive physics. These, of course, correspond to intentional and unintentional action.

The finding from the present research that reasoning about intentional and unintentional actions were uncorrelated and that the former tended to lag behind the latter throughout the preschool years, is consistent with this domain specific view of the development and nature of explaining human behavior (as well as with the dualistic sense of self, as described earlier). A goal of future research on this issue might be to devise ways to address empirically the domain-specificity versus domain-general view of explaining behavior empirically.

## CONCLUSION

The research in this dissertation begins to explore the development of the ability to explain human behavior based on the intentionality of actions, and, conversely, the ability to infer the intentionality of actions based on their explanations. As well as providing straightforward evidence that on average this double-barreled ability begins to appear near children's fifth birthday, the findings raise several other questions.

First, is the ability to infer the intentionality of action, given an explanation, more or less difficult than the reverse? Although the results from this research suggest that children find it more difficult to explain actions given their intentionality than the reverse, the research design did not provide the opportunity to address this question directly because among other reasons the two studies had different numbers of participants. However, the question could be addressed by designing the two studies to be as parallel as possible.

The findings also raise the question of whether inferring the intentionality of actions given explanations improves more throughout the preschool years than does the reverse ability. Again, the results of these studies suggest that this proposal might be true, but the evidence is not conclusive and further research is warranted.

Another question, resulting from the data indicating that children's views of intentionality almost always involve negative outcomes, concerns the role of the context in which the on purpose / not on purpose dichotomy is used in natural settings. As stated earlier, many parents said that their children use the purpose / accident distinction in order to blame a sibling or friend, or to exonerate themselves, for an act with negative consequences. It would be interesting to replicate the present studies substituting actions with negative outcomes for the ones with neutral outcomes. Such a change might lead

experimenters to find the ability addressed in this research appearing at an earlier age than seen here.

Finally, it would also be useful to come up with other ways to directly compare reasoning about intentional actions with that about unintentional actions. The data here are consistent with the proposals that children and adults are biased towards construing human action as being intentional, as well as that each type of reasoning represents a separate process. However, these are broad claims and much more research is needed to verify them.

In conclusion, the research in this dissertation, addressing children's knowledge of the relation between the intentionality of actions and their causal explanations, speaks to issues in several broad areas and themes in psychology. These include cognitive development, adult cognition, social cognition, the nature and organization of knowledge, and even the role of emotion and metaphysical beliefs in the construal of events and behavior. Clearly no single set of experiments can tackle this multi-faceted issue all at once. In fact, this issue might find a prominent place among the many fascinating and frustrating topics of debate in psychology for years to come. It will be in good company.

# Appendix

# Stories as Presented in Experiments 1 and 2

## Jungle Gym

### Experiment 1

*Intentional*: Here's Jack. Jack was playing at the jungle gym, hanging from a bar. Then he dropped down from the bar. He dropped from the bar because he wanted to go play on the swing.

*Unintentional*: Here's Henry. Henry was playing at the jungle gym too, hanging from a bar too. Then *he* dropped down from the bar. He dropped from the bar because his hands slipped off the bar.

### Experiment 2

Here's Matt. Matt was playing at the jungle gym, hanging from a bar. Then he dropped down from the bar. He meant to drop down from the bar. He did it on purpose.

## Blocks

### Experiment 1

*Intentional*: Here's Alicia. She was playing with some blocks and she made a tower. Then, she knocked over the tower. She knocked over the tower because she tripped on it.

*Unintentional*: Here's Mary. She was playing with some blocks too, and *she* made a tower. Then, she knocked over the tower. She knocked it over because she wanted to build something else.

### Experiment 2

This is Tamara. She was playing with some blocks and she made a tower. Then, she knocked over the tower. She didn't mean to knock over the tower; she didn't do it on purpose.

## Balloon

### Experiment 1

*Intentional:* This is Sally. She was at a birthday party, and someone gave her a balloon. Then, she popped the balloon. She popped the balloon because she wanted to hear the popping sound.

*Unintentional:* This is Theresa. She was at a birthday party, and someone gave her a balloon. Then, she popped the balloon. She popped the balloon because she dropped it on something sharp.

### Experiment 2

This is Casey. She was at a birthday party, and someone gave her a balloon. Then, she popped the balloon. She meant to pop the balloon. She did it on purpose.

## Crumbs

### Experiment 1

*Intentional:* This is Max. He was outside and he had cracker crumbs in his hand. Then, he dropped the crumbs onto the ground. He dropped the crumbs because he thought the birds would like them.

*Unintentional:* This is George. He was outside and he had cracker crumbs in his hand. Then, he dropped the crumbs onto the ground. He dropped the crumbs because they fell out of his hand.

### Experiment 2

This is Matt. He was outside and he had cracker crumbs in his hand. Then, he dropped the crumbs onto the ground. He didn't mean to drop the crumbs; he didn't do it on purpose.

## Jump

### Experiment 1

*Intentional:* This is Joanne. She was walking in the woods at night, and she jumped. She jumped because she wanted to get over a log.

*Unintentional:* This is Lisa. She was walking in the woods at night, and she jumped. She jumped because something scared her.

### Experiment 2

This is Hillary. She was walking in the woods at night, and she jumped. She meant to jump; she did it on purpose.

# References


Astington, J. W. (1991). Intention in the child's theory of mind. In D. Frye & C. Moore (Eds.), *Children's theories of mind: mental state and social understanding* (pp. 157-172). Hillsdale, NJ: Lawrence Erlbaum.

Astington, J. W. (1999). The language of intention: three ways of doing it. In P. Zelazo, J. Astington & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self control* (pp. 295-316). Mahway, New Jersey: Lawrence Erlbaum.

Astington, J. W. (2001). The paradox of intention. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 85-103). Cambridge, MA: MIT Press.

Baird, J. A. & Moses, L. J. (2001). Do preschoolers appreciate that identical actions may be motivated by different intentions? *Journal of Cognition and Development, 2,* 413-448.

Bartsch, K. & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60,* 946-964.

Borges, J. L. (1964). "Funes the Memorious." In *Labryrinths: Selected Stories and Other Writings.* New York: New Directions.

Browne, C. A. & Woolley, J. D. (2004). Preschoolers' magical explanations for violations of physical, social, and mental laws. *Journal of Cognition and Development, 5,* 239-260.

Buss, A. R. (1978). Causes and reasons in attribution theory: A conceptual critique. *Journal of Personality and Social Psychology, 36,* 1311-1321.

Buss, A. R. (1979). On the relationship between causes and reasons. *Journal of Personality and Social Psychology, 37,* 1458-1461.

Callanan, M. A. & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development, 7,* 213-233.

Carey, S. & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman, *Mapping the mind: Domain specificity in cognition and culture* (pp. 169-200). Cambridge, UK: Cambridge University Press.

Carpenter, M., Akhtar, A, & Tomasello, M. (2000). Fourteen- through eighteen-month-old infants differentially imitate intentional and accidental actions. In D. Muir, & A. Slater (Eds.), *Infant development: The essential readings. Essential readings in development psychology* (pp. 295-318).

Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy, 60,* 685-700.

Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes.* Cambridge, MA: MIT Press.

Feinfield, K. A., Lee, P. P., Flavell, E. R., Green, F. L., & Flavell, J. H. (1999). Young children's understanding of intention. *Cognitive Development, 14,* 463-486.

Gelman, S. A. & Kremer, K. E. (1991). Understanding natural cause: Children's explanations of how objects and their properties originate. *Child Development, 62,* 396-414.

Harre, R. & Secord, P. (1972). *The explanation of social behaviour.* Totowa, NJ: Rowman and Littlefield.

Harvey, J. & Tucker, J. (1979). On problems with the cause-reason distinction in attribution theory. *Journal of Personality and Social Psychology, 37,* 1441-1446.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: John Wiley & Sons.

Hickling, A. K. & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology, 37,* 668-683.

Hirschfeld & Gelman (1994). *Mapping the mind: Domain specificity in cognition and culture.* London: Cambridge University Press.

Hood, L. & Bloom, L. (1979). What, when, and how about why: A longitudinal study of early expressions of causality. *Monographs of the Society for Research in Child Development, 44,* 1-47.

Jones, E. E. & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219-266). New York: Academic.

Joseph, R. M. (1998). Intention and knowledge in preschoolers' conception of pretend. *Child Development, 69,* 966-980.

Joseph, R. M. & Tager-Flusberg, H. (1999). Preschool children's understanding of the desire and knowledge constraints on intended action. *British Journal of Developmental Psychology, 17,* 221-243.

Kalish, C. (1998). Reasons and causes: Children's understanding of conformity to social rules and physical laws. *Child Development, 69,* 706-720.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15, pp. 192-240).

Kruglanski, A.W. (1979). Causal explanation, teleological explanation: On radical particularism in attribution theory. *Journal of Personality and Social Psychology, 37,* 1447-1457.

Lalljee, M., Watson, M. & White, P. (1983). Some aspects of the explanations of young children. In J. Jaspars, F. Fincham, & M. Hewstone (Eds.), *Attribution theory and research: conceptual, developmental and social dimensions* (pp. 165-192). London: Academic Press.

Lennon, K. (1990). *Explaining Human Action.* London: Gerald Duckworth & Co.

Lillard, A. S. (1998). Wanting to be it: Children's understanding of intentions underlying pretense. *Child Development, 69,* 981-993.

Locke, D. & Pennington, D. (1982). Reasons and other causes: Their role in attribution processes. *Journal of Personality and Social Psychology, 42,* 212-223.

Lloyd, G. (1995) Ancient Greek concepts of causation in comparativist perspective. In D. Sperber, D. Premack, & J. Premack, *Causal cognition: A multidisciplinary debate* (pp. 536-556). Oxford: Clarendon Press.

Malle B. F. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology, 33,* 101-121.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review, 3,* 23-48.

Malle, B. F., Moses, L. J., & Baldwin, D. A. (2001). Introduction: The significance of intentionality. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 1-24). Cambridge, MA: MIT Press.

McCabe, A. & Peterson, C. (2001). A comparison of adults' versus children's spontaneous use of *because* and *so*. *Journal of Genetic Psychology, 149,* 257-268.

Moses, L. J. (1993). Young children's understanding of belief constraints on intention. *Cognitive Development, 8,* 1-25.

Russell, J., Hill, E. L., Franco, F. (2001). The role of belief veracity in understanding intentions-in-action: Preschool children's performance on the transparent intentions task. *Cognitive Development, 16,* 779-792.

Schueler (2001). Action explanation: Causes and purposes. In B. F. Malle, L. G. Moses, & D. A. Baldwin (Eds.), *Intentions and Intentionality* (pp. 251-264). Cambridge, MA: MIT Press.

Schult, C. A. (2002). Children's understanding of the distinction between intentions and desires. *Child Development, 73,* 1727-1747.

Schult, C. A. & Wellman, H. M. (1997). Explaining human movements and actions: Children's understanding of the limits of psychological explanation. *Cognition, 62,* 291-324.

Schwitzgebel, E. (1999). Gradual belief change in children. *Human Development, 42,* 283-296.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind.* Cambridge: Cambridge University Press.

Shaver, K. G. (1975). *An introduction to attribution processes.* Cambridge, MA: Winthrop.

Shultz, T. R. & Wells, D. (1985). Judging the intentionality of action-outcomes. *Developmental Psychology, 21,* 83-89.

Shultz, T. R., Wells, D. & Sarda, M. (1980). Development of the ability to distinguish intended actions from mistakes, reflexes, and passive movements. *Journal of Social and Clinical psychology, 19,* 301-310.

Smith (1978). Cognizing the behavior stream: The recognition of intentional action. *Child Development, 49,* 736-743.

Springer, K. & Keil, F. C. (1991). Early differentiation of causal mechanisms appropriate to biological and nonbiological kinds. *Child Development, 62,* 767-781.

Taylor, C. (1964). *The explanation of behavior.* London: Routledge.

Wellman, H. & Phillips, A. (2001). Developing intentional understandings. In B. F. Malle, L .J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 125-148). Cambridge, MA: MIT Press.

White, P. A. (1991). Ambiguity in the internal/external distinction in causal attribution. *Journal of Experimental Social Psychology, 27,* 259-270.

Wilson, R.A. & Keil, F.C. (2000). The shadows and shallows of explanation. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 87-114).

Woodfield, A. (1976). *Teleology.* Cambridge: Cambridge University Press.

Woolley, J. D., Browne, C. A., & Boerger, E. (In Revision). *Journal of Cognition and Development.*

# Vita

Cheryl Annette Browne was born June 1, 1973 in Buenos Aires, Argentina to Frances Rose Browne and Joseph Neil Browne. They moved to Florida, to England, and then to Chatham, New Jersey. Cheryl graduated from Chatham High School in 1991. She attended Yale University, obtaining a B.A. in psychology in 1995. She enrolled in the University of Texas at Austin graduate program in Psychology. She and Philip Matthew Bronstad were married in April 2001.

Permanent address: 59A Bristol Rd.
Medford, MA 02155

This dissertation was typed by the author.